Monte Carlo simulation of quantum statistical lattice models

Raedt, Hans De; Lagendijk, Ad

Published in:
Physics Reports

DOI:
10.1016/0370-1573(85)90044-4

Publication date:
1985

MONTE CARLO SIMULATION OF QUANTUM STATISTICAL LATTICE MODELS

Hans De RAEDT

Max-Planck-Institut für Physik und Astrophysik – Werner-Heisenberg-Institut für Physik – P.O. Box 40 12 12, Munich, Fed. Rep. Germany and Physics Department, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium

and

Ad LAGENDIJK

Natuurkundig Laboratorium, University of Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands

NORTH-HOLLAND–AMSTERDAM

PHYSICS REPORTS (Review Section of Physics Letters) 127, No. 4 (1985) 233–307. North-Holland, Amsterdam

MONTE CARLO SIMULATION OF QUANTUM STATISTICAL LATTICE MODELS

Hans De RAEDT

*Max-Planck-Institut für Physik und Astrophysik – Werner-Heisenberg-Institut für Physik – P.O. Box 40 12 12, Munich, Fed. Rep. Germany and Physics Department, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium**

and

Ad LAGENDIJK

Natuurkundig Laboratorium, University of Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands

Received March 1985

Contents:

1. General introduction 235
2. Theory 237
   2.1. Path-sum representations 237
   2.2. Small systems 241
   2.3. Monte Carlo method 242
   2.4. Improvement of the method 244
   2.5. Summary 245
3. An illustrative example 246
4. Fermions in one dimension 250
   4.1. Path integral 252
   4.2. Checkerboard representation 254
   4.3. Real-space decomposition 257
   4.4. Short chains 258
   4.5. Monte Carlo simulations 261
   4.6. Applications 266
5. Spin-1/2 models 268
   5.1. One-dimensional model 269
   5.2. Two-dimensional spin-1/2 XY model 275
   5.3. Handscomb's method: a short introduction 284
6. Systems coupled to bosons 285
   6.1. General path-integral method 285
   6.2. Electron–phonon systems 288
7. Fermions and bosons in 2 and 3 dimensions 299
8. Concluding remarks 303
References 304
Note added in proof 307

* Permanent address.

Single orders for this issue

PHYSICS REPORTS (Review Section of Physics Letters) 127, No. 4 (1985) 233–307.

Copies of this issue may be obtained at the price given below. All orders should be sent directly to the Publisher. Orders must be accompanied by check.

Single issue price Dfl. 48.00, postage included.

Abstract:

In this article we review recent developments in computational methods for quantum statistical lattice problems. We begin by giving the necessary mathematical basis, the generalized Trotter formula, and discuss the computational tools, exact summations and Monte Carlo simulation, that will be used to examine explicit examples. To illustrate the general strategy, the method is applied to an analytically solvable, non-trivial, model: the one-dimensional Ising model in a transverse field. Next it is shown how the generalized Trotter formula most naturally leads to different path-integral representations of the partition function by considering one-dimensional fermion lattice models. We show how to analyze the different representations and discuss Monte Carlo simulation results for one-dimensional fermions. Then Monte Carlo work on one- and two-dimensional spin-1/2 models based upon the Trotter formula approach is reviewed and the more dedicated Handscomb Monte Carlo method is discussed. We consider electron-phonon models and discuss Monte Carlo simulation data on the Molecular Crystal Model in one, two and three dimensions and related one-dimensional polaron models. Exact numerical results are presented for free fermions and free bosons in the canonical ensemble. We address the main problem of Monte Carlo simulations of fermions in more than one dimension: the cancellation of large contributions. Free bosons on a lattice are compared with bosons in a box and the effects of finite size on Bose–Einstein condensation are discussed.

1. General introduction

In statistical mechanics one wants to calculate thermodynamic properties of many-body systems starting from a microscopic point of view. Many of these systems cannot be (or at least have not been) solved exactly, and indeed a whole industry of approximation techniques has been developed. Many of the approximation schemes are based on perturbation theory, that is to say an expansion is set up in some physical "smallness" parameter. Examples of these approaches are high- and low-temperature expansions, small-coupling-constant expansions etc. (Mahan, 1981). Only a few systematic methods exist which are essentially non-perturbative, under which we classify Monte Carlo calculations and, to some extent, renormalization group schemes. In classical statistical mechanics the application of Monte Carlo is straightforward (Binder, 1979). Configurations in configuration space are sampled with an importance-sampling technique (often the "Metropolis" method) using the classical Boltzmann probability measure. For practical reasons the number of degrees of freedom has to be limited to $\sim 10^7$ but in many respects these small systems behave already like the infinite system. Indeed Monte Carlo calculations of classical systems have been, and still are, extremely successful in adding to our knowledge of the behavior of these methods.

The success of Monte Carlo methods for classical systems raises the question of the applicability of the Monte Carlo technique to quantum statistical mechanics. This generalization is not without any problems however: in the first place a major problem is that matrix elements of *exponents of operators* are required and in the second place the possibility of the occurrence of zero or negative matrix elements precludes straightforward application of importance sampling techniques. Let us focus on the partition function, $Z = \mathrm{Tr}\, \mathrm{e}^{-\beta H}$, for a quantum system. In a Monte Carlo experiment one would start with a certain configuration (wave function) ψ, and one would like to evaluate $\langle \psi | \mathrm{e}^{-\beta H} | \psi \rangle$. To evaluate this matrix element one generally needs to diagonalize the Hamiltonian H, which is equivalent to requiring the exact solution. If, in the exceptional case, one can evaluate easily the trace of each term that appears in the Taylor series of $\mathrm{e}^{-\beta H}$, Handscomb's method (Handscomb, 1962, 1964) can be used. Clearly one would like to have a method of much more general applicability. It is well-known that a vehicle which describes quantum mechanics in a "classical way" is the path integral. The best known path integral is the Feynman path integral (Feynman and Hibbs, 1965; Feynman, 1972). One expects that Monte Carlo calculations can be done if one can find a useful path integral representation of the partition function. Pioneering Monte Carlo work on ^{4}He using Feynman's path integral has been done by Fosdick (Fosdick, 1963) and by Fosdick and Jordan (Fosdick and Jordan, 1966; Jordan and Fosdick, 1968).

In quantum theory there are several standard procedures to obtain path integral representations of observable quantities. We will mention three of these methods:

(i) Representations based on the Stratonovich–Hubbard identity, also referred to as "uncompleting the square". The Stratonovich–Hubbard identity has stimulated a lot of approximate analytic work. As far as we know this identity has not yet been found very expedient for numerical work. However the identity has some very appealing features, and without any doubt it will prove to be useful for numerical work on quantum systems in the future. Recently a promising start has been made by using a discrete version of the Stratonovich–Hubbard identity (Hirsch, 1983b).

(ii) Introduction of anti-commuting (Grassman) variables for problems being quadratic in fermion operators (Faddeev, 1976; Schulman, 1981). Integration of quadratic functions of Grassman variables is the basis of Monte Carlo simulations of lattice gauge theories containing dynamical fermions (Fucito et al., 1981; Hamber, 1981; Weingarten and Petcher, 1981; Marinari, Parisi and Rebbi, 1981; Lang and Nicolai, 1982). In condensed-matter problems this approach has been used to study small one-dimensional many-fermion–boson models with short-range interaction in the grand-canonical ensemble (Scalapino and Sugar, 1981).

(iii) The Trotter-formula approach, examples are path integrals and Suzuki's path summations (Suzuki, 1976a), on which most of the content of this paper is based. For spin-1/2 lattice models, Suzuki showed that by employing the Trotter formula different path-sum representations can be obtained by making use of the flexibility of partitioning the Hamiltonian (Suzuki, 1976a). Subsequently Suzuki's work was extended to lattice fermion models (Barma and Shastry, 1977, 1978). For a particle in a potential, decomposition of the Hamiltonian in potential and kinetic energy most naturally leads to the Feynman path integral. In much of our work we have used the lattice analogon of the Feynman path integral for numerical simulations of quantum lattice systems. The Santa-Barbara group (Hirsch, Scalapino, Sugar and others) has implemented the Barma-and-Shastry approach for one-dimensional fermion and one-dimensional fermion–boson problems. The Barma-and-Shastry method has been extended to allow for coupling of massless fermions to a U(1) gauge field in $1+1$ dimensions (Martin and Otto, 1982).

The variety of functional integration techniques becomes even more diverse when one considers the various resolutions of the identity, like coherent- or continuous state representations (Schulman, 1981; Klauder, 1978) that have been invented.

Monte Carlo simulation techniques for many-body quantum systems fall into two distinct categories, depending whether one wants to study ground-state properties or whether one is interested in thermal properties. In the first case one can use Monte Carlo simulations to evaluate ($T = 0$) expectation values for a chosen trial wave function (McMillan, 1965; Ceperley and Kalos, 1979) or one performs a numerical integration of the Schrödinger equation by means of the Green's function Monte Carlo method (Kalos, Levesque and Verlet, 1974; Ceperley and Kalos, 1979; Kalos, 1984). These $T = 0$ Monte Carlo techniques have been applied very fruitfully to models with continuous degrees of freedom. The Green's function Monte Carlo method has been generalized to treat quantum problems at $T > 0$ and has been applied to a system of two particles interacting by a hard sphere potential (Whitlock and Kalos, 1979).

In this review we will focus on the other methods for $T > 0$ in which one usually starts from a convergent sequence of approximations for the thermal quantities. By construction this type of approach becomes useless if the temperature approaches zero. In this respect the techniques reviewed in this paper are complementary to the $T = 0$ Monte Carlo methods mentioned previously. Also we will focus on models defined on a lattice although much of what will be discussed can be transferred to related continuum-space systems with little effort (see for instance Kalos, 1984).

This review is meant as a self-contained introduction to this rapidly developing field. We hope that it will stimulate some readers to start their own calculations, because there is still so much to do. The examples we will discuss in detail are drawn mainly from our own work, the principal reason being that we understand them best. Of course we point out contributions from other workers in the specific examples we have treated. It is a review on computational techniques for quantum statistical problems. As such we find it inappropriate to discuss in detail the physics of the various models to which the method is applied because this would lead to a very incoherent review.

2. Theory

In this chapter we provide the necessary theoretical background that is needed to understand the method that will be used to calculate the thermal properties of various lattice models. Thermodynamic functions, such as energy and specific heat, can be derived from the partition function

$$Z \equiv \mathrm{Tr}\, \mathrm{e}^{-\beta H}, \tag{2.1a}$$

$$= \sum_i < \psi_i | \mathrm{e}^{-\beta H} | \psi_i \rangle, \tag{2.1b}$$

$$= \sum_\alpha w_\alpha \langle \phi_\alpha | \mathrm{e}^{-\beta H} | \phi_\alpha \rangle, \tag{2.1c}$$

whereby Tr means that we have to calculate the trace of the operator $\mathrm{e}^{-\beta H}$, H is the Hamiltonian, $\beta = 1/T$ is the inverse of the temperature T and w_α is the weight assigned to the state ϕ_α. Throughout this paper we set $k_\mathrm{B} = \hbar = 1$ and also take the lattice spacing to be one. The trace in (2.1a) can be either a sum over an orthonormal complete set of states

$$1 = \sum_i |\psi_i\rangle \langle \psi_i|, \tag{2.2a}$$

or a sum over an over-complete set of states

$$1 = \sum_\alpha w_\alpha |\phi_\alpha\rangle \langle \phi_\alpha|. \tag{2.2b}$$

Usually it is impossible to solve the eigenvalue problem for the Hamiltonian of the interacting many-body system. Then there is no general method to evaluate rigorously matrix elements such as $\langle \psi_i | \mathrm{e}^{-\beta H} | \psi_i \rangle$.

2.1. Path-sum representations

In all cases that we know of, the Hamiltonian of the interacting many-body system is a sum of several operators that can be diagonalized separately. This suggests that one should try to construct systematic approximations to $\mathrm{e}^{-\beta H}$ by using the knowledge of the spectrum of the operators that contribute to H. The Trotter formula (Trotter, 1959) and its generalizations (Suzuki, 1976b, 1977, 1984a,b) provide the

necessary mathematical justification for the construction of the approximants. It can be shown that if $H = \sum_{l=1}^{p} H_l$ we have (Suzuki, 1984a, b)

$$\|\hat{\rho} - \hat{\rho}_m\| \leq \frac{\beta^2}{2m} \sum_{l>l'=1}^{p} \|[H_l, H_{l'}]\| \exp\left[\beta \sum_{l=1}^{p} \|H_l\|\right], \tag{2.3a}$$

where

$$\hat{\rho}_m \equiv \left[\exp\left(\frac{-\beta H_1}{m}\right) \cdots \exp\left(\frac{-\beta H_p}{m}\right)\right]^m. \tag{2.3b}$$

We will call $\hat{\rho}_m$ the mth approximant to the non-normalized density matrix $\hat{\rho} = e^{-\beta H}$. From (2.3a) it follows immediately that

$$\hat{\rho} \equiv e^{-\beta H} = \lim_{m \to \infty} \hat{\rho}_m . \tag{2.3c}$$

In proving the bound (2.3a) one only has to assume that the norm ($\|\cdot\|$) of the operators H_l exists, i.e. each H_l must be bounded. Recall that a linear operator on a finite-dimensional Hilbert space is bounded and that its norm may be defined as the largest (in absolute value) eigenvalue. Hamiltonians of finite fermion- or spin-1/2 models are bounded operators.

As $\hat{\rho}_m$ is itself a product of exponential operators, we are facing the same fundamental problem as with the evaluation of $\langle \psi_i | e^{-\beta H} | \psi_i \rangle$. So far no additional assumption about the algebra of the operators H_l has been made. If all H_l's commute with each other we have the fortunate case that we can also diagonalize H itself but this is the exception rather than the rule. For some problems the operators H_l and additional well-chosen operators form a Lie algebra. Then considerable analytic progress in handling exponential operators can be made (Wilcox, 1967; Witschel, 1975; Delbourgo and Jarvis, 1983).

For many-body systems the algebra of the H_l's is often too complicated to be useful for practical applications and the only way to make progress is to use the generalized Trotter formula to approximate the partition function

$$Z = \mathrm{Tr}\, e^{-\beta H} = \lim_{m \to \infty} Z_m , \tag{2.4a}$$

by

$$Z_m = \mathrm{Tr}\, \hat{\rho}_m \tag{2.4b}$$

and to try to evaluate the trace over all states directly. Note that if all H_l are bounded (2.3) guarantees that Z_m will converge to Z if $m \to \infty$. In chapter 6 we will use the generalized Trotter formula for unbounded operators but even then it can also be shown that approximation (2.4) is meaningful (Reed and Simon, 1972; Glimm and Yaffe, 1981).

The most important decision that has to be made before actual calculations can be started is to select

the H_l's and the representation for the states in which we want to evaluate the trace. In general, there is no unique decomposition of the Hamiltonian and there is also some flexibility in choosing a representation. For definiteness we now suppose that we have chosen a certain orthonormal complete set $\{|\psi_i\rangle\}$ and insert resolutions of the identity (2.2a) between all the exponential operators appearing in the expression for Z_m. We obtain

$$Z_m = \mathrm{Tr}\,\hat{\rho}_m = \sum_{\{i_{j,l}\}} \prod_{j=1}^{m} \prod_{l=1}^{p} \langle \psi_{i_{j,l}} | \exp(-\beta H_p/m) | \psi_{i_{j,l+1}} \rangle , \tag{2.5}$$

where $|\psi_{i_{j,p+1}}\rangle = |\psi_{i_{j+1,1}}\rangle$ and $|\psi_{i_{m,p+1}}\rangle = |\psi_{i_{1,1}}\rangle$. Since we have assumed that the H_l's and ψ's have been chosen judiciously, we may suppose that each matrix element appearing in (2.5) is known analytically. Although we have approximated a positive definite operator ($e^{-\beta H}$) by another positive definite operator ($\hat{\rho}_m$) there is no reason to expect that all terms in (2.5) are positive. As a matter of fact, for many lattice models of interest there are contributions of zero or even negative weight. We will call configurations with zero weight "forbidden". Thus the remaining problem is to perform the sum over all "allowed" configurations of the system. In the rest of this chapter we will disregard special cases for which the sum (2.5) can be carried out analytically but we will concentrate on the problem of evaluating sums like (2.5) numerically.

We may interpret (2.5) as follows. We start from a configuration $|\psi_{i_{1,1}}\rangle$ and let the system propagate (in imaginary time) to a state $|\psi_{i_{1,2}}\rangle$ by applying the imaginary-time evolution operator $e^{-\beta H_1/m}$. Then the system propagates through the operator $e^{-\beta H_2/m}$ and after mp steps it arrives back at the original state $|\psi_{i_{1,1}}\rangle$. The sum in (2.5) is the sum over all possible trajectories of the system and therefore we will call it a path-summation formula. In the limit $m \to \infty$ (2.5) is formally related to the path-integral representation of the partition function of a quantum system (Feynman and Hibbs, 1965; Feynman, 1972), and as a matter of fact is identical to the Feynman path integral in certain cases.

To investigate the physical properties of a model system, we want to know the thermodynamic functions and correlations functions. We define the approximants to the free energy, energy and specific heat by

$$F_m \equiv -\frac{1}{\beta} \ln Z_m , \tag{2.6a}$$

$$E_m \equiv -\frac{1}{Z_m} \frac{\partial Z_m}{\partial \beta} , \tag{2.6b}$$

$$C_m \equiv -\beta^2 \frac{\partial E_m}{\partial \beta} . \tag{2.6c}$$

The thermal expectation value of an arbitrary operator A,

$$\langle A \rangle \equiv \mathrm{Tr}\, A\hat{\rho}/\mathrm{Tr}\,\hat{\rho} , \tag{2.7a}$$

is approximated by

$$\langle A \rangle_m \equiv \mathrm{Tr}\, A\hat{\rho}_m/\mathrm{Tr}\,\hat{\rho}_m . \tag{2.7b}$$

Note that because of (2.7b), (2.6b, c) are not the only possible definitions for the approximants of the energy and specific heat. For instance let us assume that $H = H_1 + H_2 + H_3$. Then definition (2.6b) yields

$$E_m = Z_m^{-1}(\mathrm{Tr}\, H_1[\exp(-\beta H_1/m)\exp(-\beta H_2/m)\exp(-\beta H_3/m)]^m + \mathrm{Tr}\, H_2[\exp(-\beta H_2/m)\exp(-\beta H_3/m)\exp(-\beta H_1/m)]^m + \mathrm{Tr}\, H_3[\exp(-\beta H_3/m)\exp(-\beta H_1/m)\exp(-\beta H_2/m)]^m), \tag{2.8a}$$

whereas (2.7b) gives

$$\langle H\rangle_m = Z_m^{-1}\,\mathrm{Tr}\, H[\exp(-\beta H_1/m)\exp(-\beta H_2/m)\exp(-\beta H_3/m)]^m . \tag{2.8b}$$

Although $A \neq B$ not necessarily implies $\mathrm{Tr}\, A \neq \mathrm{Tr}\, B$, comparison of (2.8a) and (2.8b) suggests that in general for finite m, $E_m \neq \langle H\rangle_m$ if $[H_1, H_2] \neq 0$, $[H_1, H_3] \neq 0$ and $[H_2, H_3] \neq 0$. For a simple numerical example of taking different definitions for the approximant to the specific heat see (De Raedt and De Raedt, 1983). In practice one uses (2.6b) and (2.6c) to estimate the thermal energy and the specific heat and (2.7b) to calculate quantities that cannot be expressed as derivatives of the free energy with respect to the inverse temperature β, the only reason being that it is convenient in Monte Carlo simulation work.

Another fundamental quantity of interest is the (Kubo) static susceptibility

$$\chi \equiv \int_0^\beta \mathrm{d}\lambda \langle \mathrm{e}^{\lambda H} A\, \mathrm{e}^{-\lambda H} B\rangle - \beta\langle A\rangle\langle B\rangle . \tag{2.9}$$

This susceptibility gives the linear response of $\langle B\rangle$ on a small perturbation of the form A. Assuming that A satisfies the same conditions as the H_l's we have

$$\chi \equiv \lim_{x\to 0} \frac{\partial}{\partial x} \frac{\mathrm{Tr}\, \mathrm{e}^{-\beta(H-xA)} B}{\mathrm{Tr}\, \mathrm{e}^{-\beta(H-xA)}}, \tag{2.10a}$$

$$\chi = \lim_{m\to\infty} \chi_m , \tag{2.10b}$$

$$\chi_m = \frac{\beta}{m} \sum_{j=1}^{m} \frac{\mathrm{Tr}\, \hat{\rho}_j A \hat{\rho}_{m-j} B}{\mathrm{Tr}\, \hat{\rho}_m} - \beta\langle A\rangle_m \langle B\rangle_m . \tag{2.10c}$$

It is of interest to know how F_m converges to F. We do not know of any rigorous solution of the general problem. In the special and also important case where $H = H_1 + H_2$ a partial answer can be given. One can show that (Golden, 1965; Symanzik, 1965; Thompson, 1965; Lieb and Thirring, 1976)

$$Z \leq Z_m , \tag{2.11a}$$

and

$$Z \leq Z_{k_{m+1}} \leq Z_{k_m} ; \qquad k_m = 2^m , \tag{2.11b}$$

or equivalently

$$F_{k_m} \leq F_{k_{m+1}} \leq F. \tag{2.11c}$$

Thus, if we partition the Hamiltonian into two parts, calculation of the mth approximant to the free energy will yield a lower bound to the exact free energy. Recall that the conventional variational principle gives upper bounds for F (Feynman, 1972). For all models and partitionings that we have studied by means of numerically exact path-summations, we always observed that the approximations to the free energy converge monotonically from below.

2.2. Small systems

An important advantage of working with lattice models instead of continuum models is that in the former case it is possible to perform (numerically) exact calculations of the static and dynamic properties of small systems whereas in the latter this cannot be done (unless there is no interaction: see chapter 7). Indeed, if we denote the number of states of the lattice model by s, the Hamiltonian can be represented by an $s \times s$ matrix and with present day computers and standard mathematical library procedures it is really straightforward to find all eigenvalues and eigenvectors as long as s is not too large. Recall that diagonalization of an $s \times s$ matrix takes of the order of s^3 operations (Wilkinson, 1965). How large s can be made in practice mainly depends on the computational resources that one has access to but typically $s \leq 2^{10}$. On the other hand, it is clear that one may attempt to generate and sum all terms that appear in (2.5). The number of contributions will be of order s^m and therefore it will be feasible to sum all terms if s and m are not too large. Obviously, summing over all contributions will require a computation time that grows exponentially with m and therefore this brute-force method is only useful for $m \leq 8$. However, bearing in mind that in this subsection we concentrate on the calculation of the properties of small systems and *not* on finding algorithms to calculate the properties of large systems, we may as well attack the path-summation problem with the same technique that we used to obtain the exact results.

We now discuss this approach only for the simplest case where the Hamiltonian is partitioned into two parts since it is straightforward to repeat all steps for any other split-up of the Hamiltonian. If $H = H_1 + H_2$ we have $Z_m = \mathrm{Tr}\, \mathscr{A}^m$ where $\mathscr{A} = \exp(-\beta H_1/m)\exp(-\beta H_2/m)$ is an $s \times s$ matrix. Calculating the mth power of $\mathscr{A}$ is trivial if the eigenvalues of $\mathscr{A}$ are known. Since it has been assumed that the model system is not too large, finding these eigenvalues is no problem. Note that in this simple example $\mathscr{A}$ could have been brought in a symmetric form ($\mathscr{A} = \exp(-\beta H_2/2m)\exp(-\beta H_1/m)\exp(-\beta H_2/2m)$) but in general $\mathscr{A}$ will not be a symmetric matrix (see for instance section 4.3) and therefore it should be taken into account that there will be left- and right-hand eigenvectors (Wilkinson, 1965). If $\mathscr{R}$ denotes the matrix of right-hand eigenvectors and $\mathscr{E}$ is the corresponding diagonal matrix of eigenvalues $(e_1 \cdots e_s)$ we have $\mathscr{R}^{-1}\mathscr{A}\mathscr{R} = \mathscr{E}$ and $Z_m = \Sigma_{n=1}^{s} e_n^m$. Thus, for any value of m the calculation of the approximation Z_m takes roughly the same time as the diagonalization of the full Hamiltonian H. Obtaining the approximate energy and specific heat without employing numerical differentation schemes is straightforward in principle but requires an additional and substantial amount of extra operations. We have

$$E_m = -\frac{m}{Z_m}\mathrm{Tr}\frac{\partial \mathscr{A}}{\partial \beta}\mathscr{A}^{m-1}, \tag{2.12a}$$

$$= -\frac{m}{Z_m} \operatorname{Tr} \mathscr{R}^{-1} \frac{\partial \mathscr{A}}{\partial \beta} \mathscr{R} \mathscr{E}^{m-1} . \tag{2.12b}$$

Thus, if we keep all model parameters and the temperature fixed the calculation of the denominator in (2.12b) requires of the order of s^4 operators. For the specific heat we find

$$C_m = \beta^2 \left[\frac{m}{Z_m} \left(\operatorname{Tr} \frac{\partial^2 \mathscr{A}}{\partial \beta^2} \mathscr{A}^{m-1} + \sum_{\lambda=1}^{m-1} \operatorname{Tr} \frac{\partial \mathscr{A}}{\partial \beta} \mathscr{A}^{\lambda-1} \frac{\partial \mathscr{A}}{\partial \beta} \mathscr{A}^{m-\lambda-1} \right) - E_m^2 \right] . \tag{2.13}$$

As it is our intention to use this approach to study the convergence of the approximations when m gets large, it is obvious that the sum over λ in the first term of (2.13) would make this scheme inefficient. Therefore we will approximate (2.13) by using the fact that the error on C_m resulting from commuting $\mathscr{A}$ and $\partial \mathscr{A} / \partial \beta$ is of higher order in β/m and we obtain

$$C_m = \beta^2 \left(\frac{m}{Z_m} \left[\operatorname{Tr} \frac{\partial^2 \mathscr{A}}{\partial \beta^2} \mathscr{A}^{m-1} + (m-1) \operatorname{Tr} \left(\frac{\partial \mathscr{A}}{\partial \beta} \right)^2 \mathscr{A}^{m-1} \right] - E_m^2 \right) . \tag{2.14}$$

Numerical evaluation of (2.14) also takes about s^4 operations. Note that (2.14) is identical to (2.13) if $m \leq 2$.

We now briefly summarize the salient features of the various techniques that can be used to study small lattice systems. Diagonalization of the full Hamiltonian provides exact numerical reference data. Exact summations over all the configurations that contribute to the various approximations for Z_m yield exact numerical results, require a computation time that grows very rapidly with m and give information about the number of "forbidden" configurations. Finally we can also diagonalize the approximant to $\exp(-\beta H)$ directly and obtain results for any value of m.

2.3. Monte Carlo method

If the number of states in (2.5) becomes so large that it would take too much computing time to add all contributions, we have to look for a different, more crude, technique that samples the most *important* contributions. To keep the notation simple we will now write $\mathscr{S}$ for the collection of intermediate states $|\psi_{i_{j,l}}\rangle$, $j = 1, \ldots, m$; $l = 1, \ldots, p$ and rewrite (2.5) as

$$Z_m = \sum_{\mathscr{S}} g(\mathscr{S}) . \tag{2.15}$$

We may interpret $\mathscr{S}$ as a closed "path" running from the state $|\psi_{i_{1,j}}\rangle$ to $|\psi_{i_{m,l}}\rangle$ and back to $|\psi_{i_{1,p}}\rangle$, the points on the path being labeled by the "imaginary-time" index j. The path must be closed because we are calculating a trace.

In statistical physics we are not really interested in the partition function itself. We want to know physical quantities such as the energy. In general a physical quantity of the quantum system can be written as

$$\langle f \rangle = \frac{\sum_{\mathscr{S}} f(\mathscr{S}) g(\mathscr{S})}{\sum_{\mathscr{S}} g(\mathscr{S})} . \tag{2.16}$$

Thus we want to compute estimators for a physical quantity by sampling the ratio of two very difficult sums. The Metropolis Monte Carlo method (Metropolis et al., 1953) is a particularly simple but very powerful algorithm for estimating the ratio

$$\langle f\rangle = \frac{\sum_{\mathcal{S}} f(\mathcal{S})\, g(\mathcal{S})}{\sum_{\mathcal{S}} g(\mathcal{S})}; \qquad \forall \mathcal{S}: g(\mathcal{S}) > 0 . \tag{2.17}$$

The majority of Monte Carlo simulations of classical statistical mechanical systems rely on this algorithm (Binder, 1979). It goes as follows. We assume that the system is in an allowed state $\mathcal{S}_{t_1}$. We then use a (pseudo) random number to choose a trial state $\hat{\mathcal{S}}$. Of course $\hat{\mathcal{S}}$ must also be an allowed configuration. Now we calculate the ratio $\mathcal{R} = g(\hat{\mathcal{S}})/g(\mathcal{S}_{t_1})$. If $\mathcal{R} \geq r$, where $0 < r < 1$ is a (pseudo) random number distributed uniformly, we replace the old state $\mathcal{S}_{t_1}$ by $\hat{\mathcal{S}}$, i.e. $\mathcal{S}_{t_1+1} \leftarrow \hat{\mathcal{S}}$. If $\mathcal{R} < r$ we keep the system in the old state $\mathcal{S}_{t_1+1} \rightarrow \mathcal{S}_{t_1}$. One can prove that, under certain restrictions to be discussed below, the sequence of states generated by this algorithm is a Markov chain of which the limiting distribution is given by $g(\mathcal{S})/\Sigma_{\mathcal{S}} g(\mathcal{S})$ (Hammersley and Handscomb, 1964). In other words, by repeating the process described above a number of times, we will obtain states that are distributed according to $g(\mathcal{S})/\Sigma_{\mathcal{S}} g(\mathcal{S})$ and then estimate $\langle f\rangle$ from

$$\langle f\rangle \approx \frac{1}{t_s} \sum_{t=t_0}^{t_s+t_0} f(\mathcal{S}_t), \tag{2.18}$$

where t_s is the number of samples taken and t_0 is the number of steps in which the system is given time to reach "equilibrium". In practice t_0 has to be determined experimentally by comparing runs with different t_0 (Binder, 1979). As we already mentioned, the partition function itself is not that interesting and this suits us very well because it is hard to calculate the partition function by means of the Metropolis Monte Carlo method (Binder, 1979). The Monte Carlo scheme sketched above is certainly not the only way to realize the Markov process. In general any Markov chain that has $g(\mathcal{S})$ as its (unnormalized) limiting distribution will do. In principle it is sufficient (but not even necessary) to construct the algorithm such that the transition probability $W(\mathcal{S}_t \rightarrow \mathcal{S}_{t'})$ satisfies the detailed balance condition $g(\mathcal{S}_t)\, W(\mathcal{S}_t \rightarrow \mathcal{S}_{t'}) = g(\mathcal{S}_{t'})\, W(\mathcal{S}_{t'} \rightarrow \mathcal{S}_t)$ and this freedom can be exploited in particular cases. The reader interested in these aspects of the Monte Carlo method itself will find an extensive discussion in (Binder, 1979).

We now look at some aspects of this Monte Carlo method in more detail. In practical applications, the most important condition that should be fulfilled is that of ergodicity. In terms of the Markov process, this means that in selecting a trial state $\hat{\mathcal{S}}$ one has to be sure that every allowed state of the system can be reached. Thereby it is not even necessary that the probability for picking out a state is uniform. In general, and we will present ample examples in other chapters, a trial state $\hat{\mathcal{S}}$ is constructed by making a small step in the multi-dimensional space, away from the current state. Although it is obvious that one cannot reach an arbitrary state by such a move, it is already sufficient if one can show that starting from any state one can get to any other state by a suitable combination of these small elementary steps. Consequently constructing an algorithm that can generate ergodic Markov chains is not difficult, the main message being: it does not matter how we get from one state to another as long as it is possible to get there. Of course in practice it may well happen that we observe that it just takes too many steps (computing time) to go from one region in configuration space to another. However if the algorithm is well-designed and we arrive at the conclusion that for certain values of the model

parameters, the system stays in a limited region of configuration space for a long period, this may be a signal that the finite system behaves as if it undergoes some kind of phase transition. Then one also expects to find large statistical fluctuations on quantities such as the specific heat and static susceptibilities.

Ideally we would like to generate a Markov chain of states such that we may consider the states $\mathcal{S}$ as statistically independent. Then standard statistical analysis on the samples of one sequence would give us an estimate for the statistical errors on physical quantities. However in practice it often happens that there are (strong) correlations between successive states in the sequence and therefore a simple statistical analysis will underestimate the true statistical fluctuations considerably. This problem is especially pertinent for some of the simulation techniques employed for quantum lattice models (see the discussions in chapters 4 and 5). A more reliable way to estimate the statistical errors is to repeat the simulation several times, using different starting configurations or random number sequences.

2.4. Improvement of the method

Unless one is very lucky it will be impossible to evaluate rigorously sums like (2.5) even in the most simple, $m = 1$, approximation. Thus in most instances one will be forced to use numerical techniques. Then one might wonder whether one can improve the approximation scheme and the efficiency of the computational technique as a whole. There are two obvious ways to accomplish this. First one can look for a faster importance-sampling technique. Taking into account that in this direction not much advancement has been made since the first proposal (Metropolis et al., 1953), this seems a very hard proposition. The remaining possibility is to improve the Trotter formula type of approximation itself. From a theoretical point of view this problem has already been studied extensively (Wilcox, 1967; Suzuki, 1976b, 1977). For instance it can be shown that (Suzuki, 1976b)

$$\hat{\rho} = \lim_{m\to\infty} \hat{\rho}_{m,n}\,, \tag{2.19a}$$

where

$$\hat{\rho}_{m,n} \equiv [\exp(-\beta H_1/m)\exp(-\beta H_2/m)\exp(\beta^2 C_2/m^2)\cdots\exp((-\beta)^n C_n/m^n)]^m\,, \tag{2.19b}$$

whereby

$$C_2 = \tfrac{1}{2}[H_2, H_1]\,, \tag{2.19c}$$

$$C_3 = \tfrac{1}{3}[C_2, H_1 + 2H_2]\,, \tag{2.19d}$$

and the general expressions for C_n can be found in the literature (Wilcox, 1967; Suzuki, 1976b). For brevity of notation we have confined ourselves to the simplest decomposition $H = H_1 + H_2$. Generalizations of (2.19) to the general case $H = \sum_{q=1}^{p} H_q$ have been given by Suzuki (Suzuki, 1976b). For $n = 1$ we obviously recover the simplest approximant (2.3b) and it is expected that the larger n the better the accuracy of the approximation will be. For simple single-particle problems it has been shown (De Raedt and De Raedt, 1983) that the use of more complicated approximants results in a more efficient and more accurate computational method. Mathematical theorems on the convergence of improved ap-

proximants have been given by Suzuki (Suzuki, 1984a, 1984b). Whether these generalized Trotter approximations can be used for many-body problems is an open question.

2.5. Summary

In the previous sections we have given the general theory that provides the basis for a computational method for calculating the thermal properties of an interacting many-body system. The aim of this current section is to review these basic concepts in the light of practical applications and to draw the attention on potential problems and shortcomings of the approach. Along the way we will give cross references to other chapters in which the reader will find more details.

Given the model Hamiltonian, the most important step is to construct a "path-sum representation" that satisfies a number of criteria. First it is clear that by breaking up the Hamiltonian, one should try to keep the most important model symmetries intact. Failure to do so results in approximants that converge slower (see the discussion on short spin-1/2 chains, section 5.1). Although inequality (2.3a) guarantees convergence of the approximants as $m \to \infty$, this knowledge is as good as useless in practical applications. Indeed (2.3a) and (2.11) are of great theoretical importance but they do not tell us much about the accuracy of approximations such as (2.6). Suppose that the left-hand side of (2.3a) is equal to the right-hand side, then in general we would have to take extremely large values of m to get some accuracy (for many-body systems $\|H\|$ is proportional to the number of particles in the system, assuming $\|H\|$ exists). This would obviously render the approach impractical. From a theoretical point of view, we can only hope that we will have reasonable accuracy for modest values of m. Taking a more pragmatic standpoint, it is clear that we can use the numerical techniques discussed in sections 2.2 and 2.3 to investigate the dependence of the rate of convergence on the size (in real space) of the system. Simple comparison of numerical data for systems of different size and fixed m should reveal whether or not the accuracy of the approximants deteriorates with increasing system size (see sections 4.4, 5.1). For most models treated in this paper it has been found that as far as the accuracy of the approximations is concerned, the system size is not an important parameter. However it should be kept in mind that there is no mathematical explanation for this nice feature of the Trotter-formula approach and one should not take it for granted when dealing with other models. In practical applications it will not be possible to use the "path-summation" approach if m becomes very large. Therefore one would like to extrapolate the finite-m results to $m \to \infty$. Then it is essential to recognize that, because of the peculiar m-dependence of the approximants Z_m, there is no theory that supports the idea of finite-size scaling with respect to the system size (m) in "imaginary-time". As a matter of fact, there exists a highly non-trivial, exactly solvable counter example where finite-size scaling in the m-direction cannot be applied (see the discussion at the end of section 5.1). Since a sound theoretical basis for finite-size scaling analysis with respect to the m-direction is lacking, we take the point of view that it is more appropriate to use standard numerical extrapolation techniques.

We now turn to problems which may and most probably will be encountered when one develops a Monte Carlo simulation method to perform the sum over all "paths". The major problem that prevents a naive application of the Monte Carlo method, as it is used in classical statistical mechanics, to quantum statistical mechanics is that in order to generate the desired Markov chain one must have $g(\mathcal{S}) > 0$ (see (2.17)). As mentioned above, most path-sum representations suffer from the fundamental drawback that $g(\mathcal{S})$ can be zero or, in some cases, negative. In most cases that we know of, the existence of states $\mathcal{S}$ for which $g(\mathcal{S}) = 0$ reflects the presence of conservation laws. It is not sufficient to devise a scheme that picks out only those states with $g(\mathcal{S}) \neq 0$. One has to take care that this scheme

gives good coverage of the phase space and that there are no excessively strong correlations between successive states of the Markov chain (see 4.5.2). If there are states for which $g(\mathcal{S})<0$ as is most evident in the case of fermion systems, one faces a very difficult fundamental problem because negative transition probabilities would make no sense. Ignoring these states is the same as making a completely uncontrolled approximation unless there are strong theoretical arguments to justify this crucial step (see 4.5.2 for such an example). General solutions to these two difficult problems are not known but as is demonstrated in the following chapters, by making use of special properties of the models the difficulties can be circumvented. Of course it is clear that the more the method exploits particular features of a model, the more difficult it becomes to use the same technique for different problems. For instance, an approach that works for 1D models can be useless for 2D or 3D systems (see chapter 4).

3. An illustrative example

To illustrate the general principles of the Trotter approach we consider a non-trivial model for which all calculations can be carried out analytically: the 1D Ising model in a transverse field. The derivation given below is taken from the work of Suzuki (Suzuki, 1976a). The Hamiltonian is given by

$$H^{\mathrm{TI}} = H_1 + H_2\,, \tag{3.1a}$$

$$H_1 = -J\sum_{i=1}^{M}\sigma_i^z\sigma_{i+1}^z\,, \tag{3.1b}$$

$$H_2 = -h\sum_{i=1}^{M}\sigma_i^x\,, \tag{3.1c}$$

where σ_i^α, $\alpha = x, y, z$, denote the Pauli-spin matrices at site i. In this chapter we always assume periodic boundary conditions. According to the general strategy outlined above, we have to specify the partitioning of Hamiltonian (3.1), the ordering of the exponential operators and a representation for the states of the system. We decompose the Hamiltonian into two parts, as indicated in (3.1a), and then the order of the exponentials makes no difference because of the cyclic permutation property of the trace. For the representation we will choose the eigenstates of σ^z and label them by Ising spin variables $S = \pm 1$ (i.e. $\sigma^z|S\rangle = S|S\rangle$). Obviously H_1 is diagonal in this representation. It is now straightforward to find the mth approximant to the partition function of model (3.1)

$$Z_m^{\mathrm{TI}} = \mathrm{Tr}[\exp(-\beta H_1/m)\exp(-\beta H_2/m)]^m \tag{3.2a}$$

$$= \sum_{\{S_{i,k}\}}\prod_{k=1}^{m}\prod_{i=1}^{M}\exp\left(\frac{\beta J}{m}S_{i,k}S_{i,k+1}\right)\langle S_{i,k}|\exp\left(\frac{\beta h\sigma_i^x}{m}\right)|S_{i,k+1}\rangle\,. \tag{3.2b}$$

As we are taking the trace we have to impose periodic boundary conditions in the Trotter direction, i.e. $S_{i,k} = S_{i,k+m}$. Evaluation of the matrix element in (3.2) gives

$$\langle S|\mathrm{e}^{a\sigma^x}|S'\rangle = (\tfrac{1}{2}\sinh 2a)^{1/2}\exp(\tfrac{1}{2}\ln\coth a)SS'\,, \tag{3.3}$$

and substituting (3.3) in (3.2b), we immediately see that Z_m^{TI} looks like the partition function of an

anisotropic two-dimensional Ising model

$$Z_m^{\mathrm{TI}} = c_m \sum_{\{S_{i,k}\}} \exp\Big[\sum_{k=1}^{m}\sum_{i=1}^{M}\Big(K_m S_{i,k} S_{i,k+1} + \frac{\beta J}{m} S_{i,k} S_{i+1,k}\Big)\Big], \tag{3.4a}$$

$$c_m = (\tfrac{1}{2}\sinh(2\beta h/m))^{mM/2} \tag{3.4b}$$

and

$$K_m = \tfrac{1}{2}\ln\coth(\beta h/m)\,. \tag{3.4c}$$

Now we can use the rigorous solution of the finite 2D Ising model (Onsager, 1944; Kaufman, 1949) to evaluate Z_m^{TI}. We have

$$Z_m^{\mathrm{TI}} = 2^{M-1}\Big(\prod_{l=1}^{M}\cosh\frac{m\gamma_{2l}}{2} + \prod_{l=1}^{M}\sinh\frac{m\gamma_{2l}}{2} + \prod_{l=1}^{M}\cosh\frac{m\gamma_{2l+1}}{2} + \prod_{l=1}^{M}\sinh\frac{m\gamma_{2l+1}}{2}\Big), \tag{3.5a}$$

where γ_l is given by

$$\cosh\gamma_l = \cosh\frac{2\beta h}{m}\cosh\frac{2\beta J}{m} - \sinh\frac{2\beta h}{m}\sinh\frac{2\beta J}{m}\cos\frac{\pi l}{M}\,. \tag{3.5b}$$

To first order in $1/m$ we obtain

$$\gamma_l = 2\beta\,\epsilon(l)/m + \mathcal{O}(m^{-2}); \qquad \epsilon(l) = (h^2 + J^2 - 2hJ\cos(\pi l/M))^{1/2}\,, \tag{3.6a}$$

and $Z_M^{\mathrm{TI}} \equiv \lim_{m\to\infty} Z_m^{\mathrm{TI}}$ is given by

$$Z_M^{\mathrm{TI}} = 2^{M-1}\Big(\prod_{l=1}^{M}\cosh\beta\,\epsilon(2l) + \prod_{l=1}^{M}\sinh\beta\,\epsilon(2l) + \prod_{l=1}^{M}\cosh\beta\,\epsilon(2l-1) + \prod_{l=1}^{M}\sinh\beta\,\epsilon(2l-1)\Big)\,. \tag{3.6b}$$

In the thermodynamic limit the free energy per site is then given by

$$-\beta\lim_{M\to\infty} F^{\mathrm{TI}}/M = \lim_{M\to\infty}\frac{1}{M}\ln Z_M^{\mathrm{TI}} = \frac{1}{2\pi}\int_0^{2\pi}\ln\Big(2\cosh\beta\,\epsilon\Big(\frac{qM}{\pi}\Big)\Big)\,\mathrm{d}q\,, \tag{3.7}$$

which is precisely the well-known result for the 1D Ising model in a transverse field (Lieb, Schultz and Mattis, 1961; Katsura, 1962; Pfeuty, 1970).

Let us now look at the approximation Z_m^{TI} in more detail. We emphasize that we have used the rigorous result for the finite Ising model. For a genuine classical system that is sufficiently large, we could have used the expression where we only retain the largest eigenvalue of the transfer matrix but it is not difficult to show that this does not lead to (3.7). In terms of transfer-matrix terminology: in order to calculate the temperature dependent properties of a d-dimensional quantum system one needs to know *all* eigenvalues of the transfer matrix of the corresponding $(d+1)$-dimensional system (Barma and

Shastry, 1978). The fundamental reason for this being that the interaction between the spins depends, in a complicated manner, on the size of the $(d+1)$-dimensional system in the additional Trotter direction. Indeed, from (3.4) we immediately see that the anistropic couplings between spins depend on the size m of the 2D lattice system. Such size-dependent interactions do not appear in partition functions of genuine classical models. Therefore the statement that the Trotter formula maps the d-dimensional spin-1/2 quantum model onto a $(d+1)$-dimensional "classical" model should not be taken too literally. There is however a special limit in which this correspondence becomes exact, namely when $T \to 0$. If $T \to 0$ we have to let $m \to \infty$ first. Without losing generality we may put $T = \Gamma/m$ and using (3.4) obtain the ground state energy per site of the 1D Ising model in a transverse field

$$E_g^{TI} = \lim_{\beta\to\infty} \lim_{M\to\infty} \lim_{m\to\infty} \frac{1}{M} \ln F_M^{TI}$$

$$= \lim_{\Gamma\to 0} \lim_{M\to\infty} \lim_{m\to\infty} \frac{\Gamma}{mM} \ln Z^{Ising}, \tag{3.8a}$$

where

$$Z^{Ising} = \sum_{\{S_{i,k}\}} \exp\left[\sum_{k=1}^{m} \sum_{i=1}^{M} \left(\tfrac{1}{2}\ln \coth(\Gamma h)\, S_{i,k} S_{i,k+1} + \Gamma J S_{i,k} S_{i+1,k}\right)\right], \tag{3.8b}$$

is the genuine classical partition function of an anisotropic 2D Ising model.

Although no new results on quantum system (3.1) have been obtained by the use of the Trotter formula, this example shows that the relationship between the d-dimensional Ising model in a transverse field and the $(d+1)$-dimensional Ising model (extending the derivation given above to a d-dimensional case is easy) (Elliott, Pfeuty and Wood, 1970; Pfeuty and Elliott, 1971; Suzuki, 1976a) makes it possible to evaluate the partition function of the quantum system by means of classical statistical mechanics techniques, without diagonalizing the quantum Hamiltonian. It is also interesting to note that one can also go the other way around. If one starts from a two-dimensional Ising model with anisotropic couplings, constructs the transfer matrix and adopts the τ-continuum Hamiltonian approach one arrives at model (3.1) (Kogut, 1979).

The next step in our program is to investigate the convergence of the approximant Z_m^{TI} to the partition function. For the case under consideration this is fairly simple because numerical evaluation of (3.5–7) is straightforward. In table 3.1 we give typical results for the approximations to the free energy. The difference between F_M^{TI} and F^{TI} is due to finite size effects. From these numbers we may conclude that, even in the case of very low temperature (compare the ground-state energy with the exact free energy), the rate of convergence is quite good. According to inequality (2.11c) the approximations F_m^{TI} should give lower bounds to the exact free energy and table 3.1 shows that this is indeed the case. A more detailed analysis of the error $|1 - F_m^{TI}/F_M^{TI}|$ is given in fig. 3.1 where we have plotted the relative errors for various choices of the model parameters. From this figure we see that by increasing the field h and keeping all other parameters constant, the approximation becomes less accurate as long as h is not too large (for $T = 1$ and $M = 32$ this means $h/J < 3$). For large h (which for $T = 1$ and $M = 32$ means $h/J > 3$) the accuracy of the approximation starts to increase again as it should because from (3.2) we immediately see that Z_m^{TI} gives the exact answers for *any* value of m if $J = 0$ or $h = 0$. The conclusions that

Table 3.1
The approximation to the free energy per site (F_m^{TI}, see (3.5)) of the 1D Ising model in a transverse field as a function of the number of products m at low temperature $T/J = 0.50$. The field $h/J = 0.50$ and the number of sites $M = 32$. The ground-state energy per site $E_0^{TI} = -1.0635$ and the free energy per site in the thermodynamic limit $F^{TI} = -1.0798$

m	F_m^{TI}	Relative error	F_M^{TI}
1	−1.2302	0.1209	−1.0815
2	−1.1409	0.0521	−1.0815
3	−1.1117	0.0272	−1.0815
4	−1.0995	0.0163	−1.0815
5	−1.0933	0.0108	−1.0815
8	−1.0863	0.0044	−1.0815
10	−1.0846	0.0028	−1.0815
16	−1.0827	0.0011	−1.0815
32	−1.0818	0.0003	−1.0815
64	−1.0816	0.0001	−1.0815

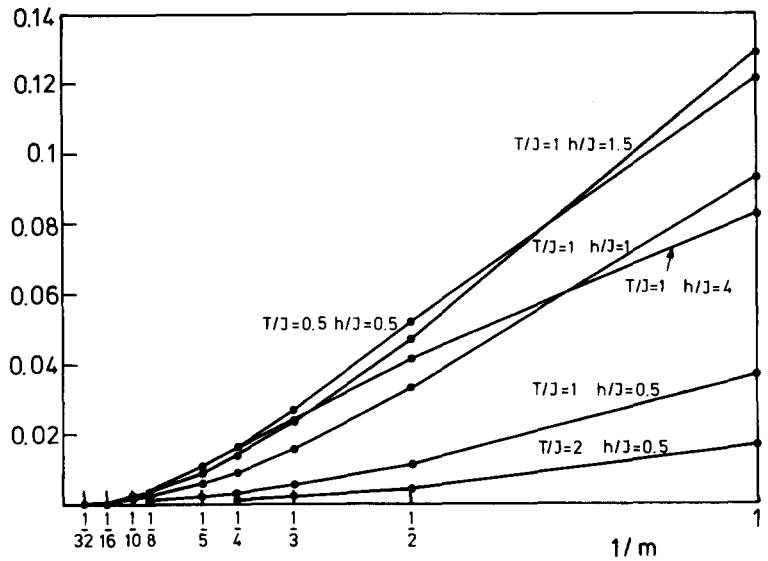

Fig. 3.1. The relative error of the approximation to the free energy of the 1D Ising model in a transverse field obtained from (3.5) with respect to the exact result for the free energy given by (3.6) for several values of the parameters of the Hamiltonian (3.1).

we draw from this is that for fixed T and m, the relative error reaches a maximum for specific values of the model parameters. The fact that there is a special point in the parameter space of the model ($h = J$) is not of any special significance for the accuracy of the Trotter approximation.

Let us now investigate what would happen if we would try to utilize the Monte Carlo method. Of course performing simulations for model (3.1) would really be a waste of computer time since, as shown above, a much more accurate analysis of the approximations can be done without the use of sampling techniques. It is well-known that one should not attempt to calculate (3.5) by means of the Metropolis Monte Carlo method (Binder, 1979) but the simulation technique can be used to obtain physical

quantities such as the energy, specific heat etc. We will confine ourselves to a discussion of the approximation to the energy which is given by

$$E_m^{\mathrm{TI}}/M = -h \coth\frac{2\beta h}{m} + h\,\mathrm{csch}\frac{2\beta h}{m}\langle S_{i,k}S_{i,k+1}\rangle_m - J\langle S_{i,k}S_{i+1,k}\rangle_m\,, \tag{3.9}$$

because this is sufficient to demonstrate a serious fundamental difficulty that may arise if one uses the Monte Carlo method to evaluate path sums. Note that as m get larger, K_m increases and consequently the spins in the Trotter direction become more aligned. On the other hand, if $m\to\infty$, $\coth(2\beta h/m)\approx \mathrm{csch}(2\beta h/m)\approx m/2\beta h$ and as the energy per site is of order one we need to have $\langle S_{i,k}S_{i,k+1}\rangle \approx 1-\mathcal{O}(m^{-1})$, $m\to\infty$. Thus for large m we have to substract two large numbers of order m whereas the final result is of order one. Because of this suspect numerical procedure we conclude that increasing m without carefully studying the statistical errors is a dangerous way to proceed. This difficulty is not encountered only in Monte Carlo simulations of model (3.1) (Wiesler, 1982) but also appears in even simpler applications (the harmonic oscillator for instance) of the Monte Carlo method in path integral calculations (Herman, Bruskin and Berne, 1982).

4. Fermions in one dimension

In this chapter we apply the ideas of chapter 2 to the one-dimensional spinless fermion lattice model

$$H = H_1 + H_2\,, \tag{4.1a}$$

$$H_1 = -t\sum_{i=1}^{M}(c_i^+ c_{i+1} + c_{i+1}^+ c_i)\,, \tag{4.1b}$$

$$H_2 = v_1\sum_{i=1}^{M} c_i^+ c_i c_{i+1}^+ c_{i+1} = v_1\sum_{i=1}^{M} n_i n_{i+1}\,. \tag{4.1c}$$

The fermion operator $c_i^+(c_i)$ creates (annihilates) a particle at site i, the operator $n_i \equiv c_i^+ c_i$ counts the number of fermions at sites i, the total number of particles being $N = \sum_{i=1}^{M} n_i$ where M is the number of sites of the chain. In our numerical work we set the hopping energy t equal to 1. The strength of the nearest-neighbor interaction is given by v_1. The density of particles will be denoted by $\rho \equiv N/M$.

Denoting spin-raising and spin-lowering operators by $\sigma_l^+ = (\sigma_l^x + \mathrm{i}\sigma_l^y)/2$ and $\sigma_l^- = (\sigma_l^x - \mathrm{i}\sigma_l^y)/2$ respectively, we can use the Jordan–Wigner transformation

$$c_l^+ = \sigma_l^+ \exp\left[\frac{\mathrm{i}\pi}{2}\sum_{p=1}^{l-1}(1+\sigma_p^z)\right], \tag{4.2a}$$

$$c_l = \sigma_l^- \exp\left[-\frac{\mathrm{i}\pi}{2}\sum_{p=1}^{l-1}(1+\sigma_p^z)\right], \tag{4.2b}$$

to rewrite spinless fermion model (4.1) as a spin-1/2 model

$$H = -\frac{t}{2}\sum_{i=1}^{M}\left(\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y - \frac{v_1}{2t}\sigma_i^z\sigma_{i+1}^z - \frac{v_1}{t}\sigma_i^z - \frac{v_1}{2t}\right). \tag{4.3}$$

If free boundary conditions are adopted for the spinless fermion model, the correspondence between (4.1) and (4.3) is exact. For periodic boundary conditions an extra term has to be added to (4.3) due to the non-local character of transformation (4.2). This term causes the sign of the terms involving $\sigma_M^\alpha\sigma_1^\alpha$, $\alpha = x, y$ to depend on the total magnetization, i.e. the number of fermions N, of the state of system. Although it might be argued that this is only a boundary effect and hence not important if $M\to\infty$ (N/M constant), it has to be kept in mind that the spectrum of fermion Hamiltonian (4.1) is not completely equivalent to the spectrum of spin Hamiltonian (4.3). In particular if *all* eigenvalues of (4.3) were known, it would be possible to infer only one half of the energy levels of (4.1) (Wolf and Zitartz, 1981). Our numerically exact results for the spectrum of spinless fermions on short rings show that all eigenvalues are at least doubly degenerate if the system is half filled ($\rho = \frac{1}{2}$).

Starting with the introduction of the Bethe ansatz (Bethe, 1931), a vast amount of rigorous results on spin model (4.3) and its generalization, the X-Y-Z model, have been obtained over the past fifty years. This knowledge gathered by combinations of analytical and numerical techniques may be considered as a major achievement of mathematical physics of many-body systems. As it has taken considerable time and effort to unravel most of the model properties (not all important questions have been answered yet (Bonner et al., 1981)) it should be no surprise that the simple looking one-dimensional model (4.1) or its spin-1/2 form (4.3) are both highly non-trivial many-body systems in which quantum effects are very important.

For the spinless fermion model (4.1) quantities of interest are thermodynamic functions and the structure factor

$$S(q) \equiv \sum_{j=1}^{M}\left(\langle n_i n_{i+j}\rangle - \langle n_i\rangle\langle n_j\rangle\right)\cos qj, \tag{4.4a}$$

$$= \langle\rho_q\rho_{-q}\rangle - \delta_{q,0}\langle\rho_q\rangle^2, \tag{4.4b}$$

which gives us detailed information about the arrangement of the particles on the chain. The behavior of $S(q)$ as a function of M and temperature tells us whether or not there will be long range order in the ground state. For $\rho = \frac{1}{2}$ the ground state of the spinless fermion model corresponds to the ground state of the antiferromagnetic spin-1/2 model. Using the relationship with the spin-1/2 chain, it has been shown that the fermion system undergoes a metal–insulator transition at $v_1 = 2t$ (Ovchinnikov, 1973). If $v_1 < 2t$ there is no energy gap in between the ground state energy and the first excited states and the system is a metal (Fowler and Puga, 1978). Then there is only "incipient" long-range order such that $S(q)$ peaks at $q = \pi$ but has a finite width. If $v_1 > 2t$ there is a gap and the ground state has long-range order which, for $v_1\to\infty$ approaches the ground state of the classical ($t = 0$) model: $|1010\ldots\rangle$. Obviously in this ground state there is long-range order and if we calculate the structure factor for this particular configuration we find $S(q) = M\delta(\pi - |q|)$. Although the gap in the energy spectrum cannot be calculated by the approach pursued here, we can study the metal–insulator transition by looking for a drastic change in the behavior of $S(q)$ as $v_1/2t$ goes from zero to infinity. For $\rho \neq \frac{1}{2}$ the system is always metallic (Fowler and Puga, 1978).

The static structure factor (4.4) is a measure for the fluctuations of the density and is closely related to the static density–density (Kubo) susceptibility

$$\chi(q) = \int_0^\beta \mathrm{d}x \, \langle \mathrm{e}^{xH} \rho_q \, \mathrm{e}^{-xH} \rho_{-q} \rangle - \beta \delta_{q,0} \langle \rho_q \rangle^2 , \tag{4.5}$$

through the fluctuation-dissipation theorem. The susceptibility (4.5) measures the response of the (Fourier transformed) density operator ρ_q to an external density perturbation and is bounded by the Bogoliubov inequality $\chi(q) \leq \beta S(q)$. Recall that in classical statistical mechanics one has $\chi(q) = \beta S(q)$. As a consequence the difference between $\chi(q)$ and $\beta S(q)$ gives some indication of the importance of quantum effects.

Since so much is known about (4.1) or (4.3) we believe any new computational technique for solving systems of interacting quantum particles should be tested on at least one of these one-dimensional models. Then there can be no uncertainty in the interpretation of the results obtained by the new technique and it should be possible to make a critical evaluation of the accuracy, efficiency and deficiencies of the new technique. There is however no guarantee that the computational method that has passed this test successfully, will also work for 2D or 3D models.

In the first part of this chapter we carry out such a program. We introduce three different representations for the partition function of model (4.1) subject to periodic boundary conditions, and follow the strategy outlined in chapter 2. Application of the path-summation approach to spin-1/2 model (4.3) and a more detailed discussion of its relationship with (4.1) is postponed to the next chapter. There we show, among other things, that all known mappings between various models follow directly from particular path-sum representations. The last part of this chapter 4 is devoted to applications of the path-summation approach to generalizations of (4.1) which are not amenable to the more conventional treatments. In particular we include more-distant interactions and also consider fermions with spin. Further extensions to models for fermions interacting with bosons are given in chapter 6.

4.1. Path integral

In this section we derive the genuine lattice form of the path integral (PI) of an interacting spinless fermion model. To obtain the path integral form we decompose the Hamiltonian in kinetic and potential energy. The corresponding Trotter approximation reads

$$Z_m^{\mathrm{PI}} = \mathrm{Tr}[\exp(-\beta H_1/m) \exp(-\beta H_2/m)]^m . \tag{4.6}$$

Obviously this approximation will reproduce the exact partition function for any value of m if we take the classical limit (which we define by putting $t = 0$ in (4.1)) or consider the free quantum system $v_1 = 0$. From (2.11) it follows that for finite m (4.6) will give us lower bounds to the exact free energy.

In the occupation number representation resolutions of the identity operator are given by

$$1 = \sum_{\{i_1 < \cdots < i_N\}} c_{i_1}^+ \cdots c_{i_N}^+ |0\rangle \langle 0| c_{i_N} \cdots c_{i_1} \tag{4.7a}$$

$$= \frac{1}{N!} \sum_{\{i_l\}} c_{i_1}^+ \cdots c_{i_N}^+ |0\rangle \langle 0| c_{i_N} \cdots c_{i_1} . \tag{4.7b}$$

Note that in a rather trivial way, (4.7b) is a sum over a set of over-complete states. Inserting (4.7a) between all exponents in (4.6) and using the fact that the potential energy is diagonal in the occupation number representation yields

$$Z_m^{\mathrm{PI}} = \sum_{\{i_{1,j}<\cdots<i_{N,j}\}} \prod_{j=1}^{m} \langle 0|c_{i_{N,j}} \cdots c_{i_{1,j}} \mathrm{e}^{-\beta H_1/m}\, c^+_{i_{1,j+1}} \cdots c^+_{i_{N,j+1}}|0\rangle$$

$$\times \exp\Big(-\frac{\beta v_1}{m} \sum_{\mu,\nu=1}^{N} \delta_{|i_{\mu,j}-i_{\nu,j}| \bmod M,\, 1}\Big)\,. \tag{4.8}$$

To evaluate the remaining matrix elements in (4.8) we write the many-fermion state in its Fourier representation

$$c^+_{l_1} \cdots c^+_{l_N}|0\rangle = M^{-N/2} \sum_{\{k_\mu\}} \Big(\prod_{\mu=1}^{N} \exp\frac{2\pi \mathrm{i} k_\mu l_\mu}{M}\Big) c^+_{k_1} \cdots c^+_{k_N}|0\rangle\,, \tag{4.9a}$$

where $k_\mu = 1, \ldots, M$, we use the knowledge that H_1 becomes diagonal after the Fourier transformation

$$H_1 c^+_{k_1} \cdots c^+_{k_N}|0\rangle = \Big(-2t \sum_{\mu=1}^{N} \cos\frac{2\pi k_\mu}{M}\Big) c^+_{k_1} \cdots c^+_{k_N}|0\rangle\,, \tag{4.9b}$$

and employ the identity

$$\langle 0|c_{q_N} \cdots c_{q_1} c^+_{k_1} \cdots c^+_{k_N}|0\rangle = \sum_{P} \mathrm{sign}(P)\, \delta_{k_1,q_{P1}} \cdots \delta_{k_N,q_{PN}}\,. \tag{4.9c}$$

Here P is the permutation operator acting on the set $\{1, \ldots, N\}$, $\mathrm{sign}(P) = +1$ (-1) if the parity of P is even (odd) and the sum in (4.9c) is over all possible permutations of the N particles. We obtain

$$\langle 0|c_{i_{N,j}} \cdots c_{i_{1,j}} \mathrm{e}^{-\beta H_1/m} c^+_{i_{1,j+1}} \cdots c^+_{i_{N,j+1}}|0\rangle = \sum_{P_j} \prod_{\mu=1}^{N} \mathrm{sign}(P_j)\, I\Big(\frac{2\beta t}{m}, i_{\mu,j} - i_{P_j\mu,j+1}\Big)\,, \tag{4.10}$$

where

$$I(z, l) = \frac{1}{M} \sum_{n=1}^{M} \cos\frac{2\pi l n}{M} \exp\Big(z \cos\frac{2\pi n}{M}\Big)\,, \tag{4.11}$$

is the Fourier transformed imaginary-time lattice propagator of a free particle and is closely related to the modified Bessel functions $I_l(z)$ by $I_l(z) = \lim_{M\to\infty} I(z, l)$. It also obeys the recursion relation

$$2\frac{\mathrm{d}}{\mathrm{d}z} I(z, l) = I(z, l-1) + I(z, l+1)\,. \tag{4.12}$$

Inserting (4.10) in (4.8) yields

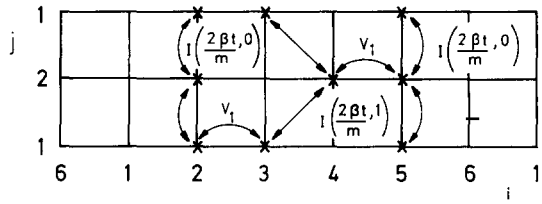

Fig. 4.1. Graphical representation of one particular contribution to Z_2^{PI} (see (4.13)) for 3 spinless fermions moving on a ring of 6 sites. Arrows indicate the various kinds of interactions.

$$Z_m^{\mathrm{PI}} = \sum_{\{i_{1,j}<\cdots<i_{N,j}\}} \sum_{\{P_j\}} \prod_{j=1}^{m} \prod_{\mu,\nu=1}^{N} \mathrm{sign}(P_j)\, I\Big(\frac{2\beta t}{m}, i_{\mu,j} - i_{P_j\mu,\,j+1}\Big) \exp\Big(-\frac{\beta v_1}{m}\,\delta_{|i_{\mu,j}-i_{\nu,j}|\bmod M,\,1}\Big)\,. \tag{4.13}$$

The restriction on the sum over $i_{\mu,j}$ and the sum over all possible permutations can be combined to simplify (4.13) further

$$Z_m^{\mathrm{PI}} = \frac{1}{N!} \sum_{\{y_{\mu,j}\}} \sum_{P} \prod_{j=1}^{m} \prod_{\mu,\nu=1}^{N} \mathrm{sign}(P)\, I\Big(\frac{2\beta t}{m}, y_{\mu,j} - y_{\mu,j+1}\Big) \exp\Big(-\frac{\beta v_1}{m}\,\delta_{|y_{\mu,j}-y_{\nu,j}|\bmod M,\,1}\Big)\,, \tag{4.14}$$

whereby each sum over the y's runs over all possible sites and $y_{\mu,m+1} \equiv y_{P\mu,1}$. Formally (4.14) is just the genuine lattice version of the Feynman path integral of an interacting fermion system (Feynman, 1972; Wiegel, 1975). We have found (4.13) more suitable for Monte Carlo work because it allows a simple implementation of the algorithm that generates permutations. For the one-dimensional model (4.1), we can interpret expression (4.13) or (4.14) in terms of a two-dimensional lattice model as is illustrated in fig. 4.1.

As explained in section 2.2, explicit expressions for the approximants to thermodynamic functions or correlation functions can be derived directly from representation (4.13) or (4.14). Formulas for the energy, specific heat and other quantities can be found in the original papers (De Raedt and Lagendijk, 1981, 1982a, 1983a).

There are several, almost trivial, extensions of this type of representation. First it is straightforward to write down the formal expression of the path integral of a 2D or 3D system. This only makes the notation more cumbersome. Second it is also possible to include more-distant interactions between the fermions since such terms are diagonal in the occupation number representation. Models with additional next-nearest neighbor interactions cannot be handled by the majority of the techniques which have been so successful in producing the important results on model (4.1) but can be treated by this particular path-summation approach without much extra effort.

4.2. Checkerboard representation

In this and the following section we explicitly use the extreme local character of model (4.1) to construct path sum representations. Following Barma and Shastry (Barma and Shastry, 1977) we write the Hamiltonian (4.1) as $H = A + B$ where

$$A = \sum_{i=1}^{M/2} H_{2i-1,2i}\,; \qquad B = \sum_{i=1}^{M/2} H_{2i,2i+1}\,, \tag{4.15}$$

and the Hamiltonian $H_{i,j}$ is given by

$$H_{i,j} = -t(c_i^+ c_j + c_j^+ c_i) + v_1 n_i n_j . \tag{4.16}$$

For simplicity we have assumed that the number of sites M is even. As a direct consequence of the local character of $H_{i,j}$ we have

$$[H_{2i-1,2i}, H_{2j-1,2j}] = [H_{2i,2i+1}, H_{2j,2j+1}] = 0 , \tag{4.17}$$

and therefore the corresponding Trotter approximation reads

$$\begin{aligned} Z_m^{\mathrm{CBD}} = \mathrm{Tr}[&\exp(-\beta H_{1,2}/m)\exp(-\beta H_{3,4}/m)\cdots\exp(-\beta H_{M-1,M}/m) \\ &\times \exp(-\beta H_{2,3}/m)\cdots\exp(-\beta H_{M-2,M-1}/m)\exp(-\beta H_{M,1}/m)]^m . \end{aligned} \tag{4.18}$$

As we will show below this approximation can be reformulated in terms of particles moving on a checkerboard lattice and therefore we call (4.18) the checkerboard decomposition (CBD). Since we partition the Hamiltonian in two parts A and B, it follows from (2.11) that (4.18) will also provide upperbounds (lowerbounds) to the exact partition function (free energy).

The two-dimensional lattice model corresponding to (4.18) is obtained by inserting the resolutions of identity (4.7a). In chapter 5 we also use the local partitioning to study spin-1/2 models. Since the formal manipulations for fermions and spins are almost identical, we can avoid duplication of a number of formulas and ultimately save a lot of extra computer programming by denoting fermion states in a spin-like language. Thus we will denote a state of the fermion system in the occupation number representation by $|S_1, \ldots, S_M\rangle$ where $S_i = 1$ (-1) if the site i is occupied (empty). It is now straightforward to obtain (De Raedt and Lagendijk, 1982b)

$$\begin{aligned} Z_m^{\mathrm{CBD}} = \sum_{\{S_{i,j}\}}{}' \sum_{\{\bar{S}_{i,j}\}}{}' \prod_{j=1}^{m} & T(S_{1,j}, S_{2,j}; \bar{S}_{1,j}, \bar{S}_{2,j}) \cdots T(S_{M-1,j}, S_{M,j}; \bar{S}_{M-1,j}, \bar{S}_{M,j}) \\ & \times T(\bar{S}_{2,j}, \bar{S}_{3,j}; S_{2,j+1}, S_{3,j+1}) \cdots T(\bar{S}_{M-2,j}, \bar{S}_{M-1,j}; S_{M-2,j+1}, S_{M-1,j+1}) \\ & \times T(\bar{S}_{M,j}, \bar{S}_{1,j}; S_{M,j+1}, S_{1,j+1})(1 - |S_{1,j} - S_{M,j}|)^{N+1} , \end{aligned} \tag{4.19}$$

where

$$\begin{aligned} T(S_i, S_j; \bar{S}_i, \bar{S}_j) &\equiv \langle S_i, S_j | \exp(-\beta H_{i,j}/m) | \bar{S}_i, \bar{S}_j \rangle \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cosh(\beta t/m) & \sinh(\beta t/m) & 0 \\ 0 & \sinh(\beta t/m) & \cosh(\beta t/m) & 0 \\ 0 & 0 & 0 & \exp(-\beta v_1/m) \end{pmatrix} ; \begin{matrix} |-1\,-1\rangle \\ |\ 1\,-1\rangle \\ |-1\ \ 1\rangle \\ |\ 1\ \ 1\rangle \end{matrix} . \end{aligned} \tag{4.20}$$

The primes on the summation signs in (4.19) indicate that the sum over the S and $\bar{S}$ variables is restricted because the total number of fermions N is fixed, i.e. $\Sigma_{i=1}^{M} S_{i,j} = \Sigma_{i=1}^{M} \bar{S}_{i,j} = 2N - M$ for all j. As we have chosen to work with periodic boundary conditions, there exists the possibility that the

operator $e^{-\beta A/m}\, e^{-\beta B/m}$ moves a particle from site 1 to site M or vice versa. Such moves destroy the order in which the fermions have been created from the vacuum state. The last factor in (4.19) is the correction term that results from reordering the fermion operators. Obviously there are only negative contributions if the number of particles is even and no minus signs would be present for free boundary conditions.

As for the path-integral representation it is straightforward to write down the checkerboard form for a 2D or 3D model (Barma and Shastry, 1977) but unfortunately the extremely nice feature of the 1D model, i.e. no negative contributions in the case of free boundary conditions, is lost in the case of a two- or three-dimensional system. Indeed it is not difficult to see that keeping the order in which fermions are created from the vacuum state consistent with a previously chosen convention, unavoidably leads to negative contributions to Z_m^{CBD}, whatever the initial choice of the boundary conditions.

From (4.19) and its graphical representation fig. 4.2 we conclude that in this formulation the number of variables is twice that of the path-integral representation (4.13). However it is possible to perform the sum over all $\bar{S}_{i,j}$ analytically by exploiting the special form of $T(S_i, S_j; \bar{S}_i, \bar{S}_j)$. The analytic elimination of degrees of freedom that is carried out here is formally the same as for the path integral. Here we sum out all $\bar{S}_{i,j}$, to obtain the path integral (4.14) the wave numbers k_μ have been summed out. We rewrite (4.20) as

$$T(S_i, S_j; \bar{S}_i, \bar{S}_j) = \delta_{S_i S_j,\, \bar{S}_i \bar{S}_j} T_{S_i \bar{S}_i}(S_i, S_i \bar{S}_i S_j)\,, \tag{4.21a}$$

where

$$T_1(S, \bar{S}) \equiv \begin{pmatrix} 1 & \cosh(\beta t/m) \\ \cosh(\beta t/m) & \exp(-\beta v_1/m) \end{pmatrix}; \quad \begin{matrix} |-1\rangle \\ |\ 1\rangle \end{matrix}, \tag{4.21b}$$

and

$$T_{-1}(S, \bar{S}) \equiv \delta_{S,\bar{S}} \sinh(\beta t/m)\,. \tag{4.21c}$$

Substituting (4.21b) in (4.19) we can sum out almost all $\bar{S}_{i,j}$ and we find

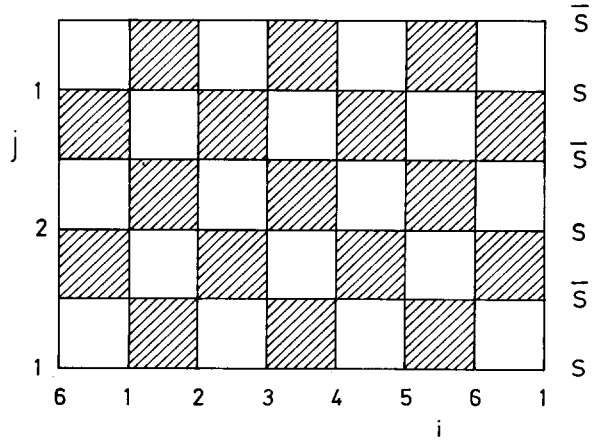

Fig. 4.2. Graphical representation of the 2D effective lattice corresponding to Z_2^{CBD} (see (4.19)) of a ring of 6 sites. Each shaded square represents a T-matrix and also determines which particles can interact with each other. Only particles that sit on the corners of the same shaded square feel each other. Note that compared with fig. 4.1 there are twice as many intermediate states to sum over. The variables $S_{i,j}$ (see (4.19)) are defined on the rows $j = 1, 2$, whereas the variables $\bar{S}_{i,j}$ are defined on the rows between the $j = 1$ and $j = 2$ rows.

$$Z_m^{\mathrm{CBD}} = \sum_{\{S_{i,j}\}} \sum_{\{\sigma_j\}} \prod_{j=1}^{m} T_{\sigma_j \phi_{1,j}}(S_{1,j}, \sigma_j \phi_{1,j} S_{2,j})$$
$$\times T_{\sigma_j \phi_{2,j}}(\sigma_j \phi_{2,j} S_{2,j}\, S_{3,j+1}) \cdots T_{\sigma_j \phi_{M-1,j}}(S_{M-1,j}, \sigma_j \phi_{M-1,j} S_{M,j})$$
$$\times T_{\sigma_j \phi_{M,j}}(\sigma_j \phi_{M,j} S_{M,j}, S_{1,j+1}) \sigma_j^{N+1} \delta_{\phi_{M,j},1}\,, \tag{4.22}$$

where

$$\phi_{l,j} \equiv \prod_{i=1}^{l} S_{i,j} S_{i,j+1}\,. \tag{4.23}$$

From expression (4.22) we conclude that the number of variables has been reduced from $2mM$ to $m(M+1)$ at the expense of introducing more complicated "string-like" variables. The presence of these strings is an indication that the effective lattice model (4.22) contains non-local interactions along the chain and in the additional Trotter direction. The extra sum over the σ's is a direct consequence of working with periodic boundary conditions. As we keep the number of fermions fixed, the constraint $\delta_{\phi_{M,j},1}$ in (4.22) is automatically satisfied but we did not omit it because we will make use of (4.22) in our treatment of spin-1/2 systems (see chapter 5) where it is not automatically satisfied.

The checkerboard partitioning has a peculiar property. From section 2.5, we know that the replacement of $\exp(A+B)$ by $\exp(A)\exp(B)$ becomes exact if $C_2 = [B, A]/2 = 0$. Specializing to the non-interacting case $v_1 = 0$ and using the Fourier-transformed form of the fermion operators we obtain $C_2 = 2\mathrm{i} \sum_{k=1}^{M} c_k^{+} c_{k+M/2} \sin(4\pi k/M)$. Obviously there are four values of k ($k = M/4,\ M/2,\ 3M/4,\ M$) for which the sine in C_2 vanishes. Therefore the checkerboard partitioning will give us the exact results for the free-fermion system if and only if the number of sites $M = 4$.

4.3. Real-space decomposition

Here we follow Suzuki's original proposal (Suzuki, 1976a) for the spin-1/2 chain and we order the exponential operators in the same way as the two-site blocks make up the chain. We call this partitioning the real-space decomposition (RSD). We follow the same procedure as in the previous section to express

$$C_m^{\mathrm{RSD}} \equiv \mathrm{Tr}[\exp(-\beta H_{1,2}/m) \cdots \exp(-\beta H_{M-1,M}/m) \exp(-\beta H_{M,1}/m)]^m\,, \tag{4.24}$$

as a path summation formula and obtain (De Raedt and Lagendijk, 1982a)

$$Z_m^{\mathrm{RSD}} = {\sum_{\{S_{i,j}\}}}' {\sum_{\{\bar{S}_{i,j}\}}}' \prod_{j=1}^{m} T(S_{1,j}, S_{2,j}; \bar{S}_{1,j}, \bar{S}_{2,j})$$
$$\times T(\bar{S}_{2,j}, S_{3,j}; S_{2,j+1}, \bar{S}_{3,j}) \cdots T(\bar{S}_{M-1,j}, S_{M,j}; S_{M-1,j+1}, \bar{S}_{M,j})$$
$$\times T(\bar{S}_{M,j}, \bar{S}_{1,j}; S_{M,j+1}, S_{1,j+1})(1 - |S_{1,j} - S_{M,j}|)^{N+1} \tag{4.25}$$
$$= {\sum_{\{S_{i,j}\}}}' \sum_{\{\sigma_j\}} \prod_{j=1}^{m} \prod_{i=1}^{M} T_{\sigma_j \phi_{i,j}}(\phi_{i-1,j} S_{i,j}, \phi_{i,j} S_{i+1,j}) \sigma_j^{N+1} \delta_{\phi_{M,j},1}\,,$$

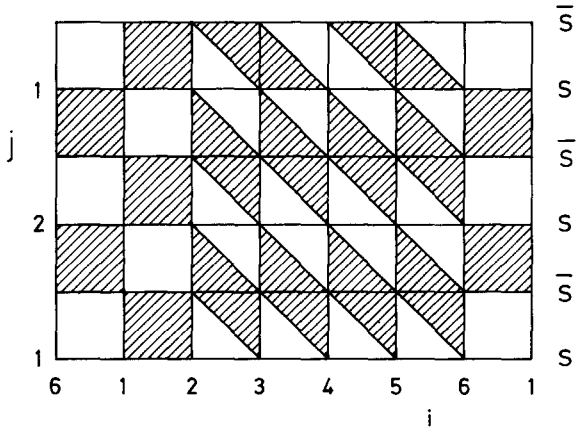

Fig. 4.3. Graphical representation of the 2D effective lattice corresponding to Z_2^{RSD} (see (4.25)) of a ring of 6 sites. Each shaded area represents a T-matrix and also determines which particles can interact with each other. Only particles that sit on the corners of the same shaded area feel each other. Remark that real-space partitioning combined with periodic boundary conditions in the physical dimension destroys the regularity of the 2D effective lattice.

where $\phi_{0,j} \equiv 1$. Comparing the graphical interpretation of (4.25) displayed in fig. 4.3 with fig. 4.2 we conclude that both lattices have similar properties except at the sites 1 and M where in the case of the real-space decomposition, the regularity of the lattice structure is destroyed due to the periodic boundary conditions. This implies that for any finite value of m, static correlation functions such as $\langle n_i n_j \rangle$ will depend on whether or not i or j is equal to 1 or M. In this sense this real-space approach breaks the translation symmetry of the original model, but in practice this is not a serious limitation because we will normally calculate translational invariant physical quantities by summing over all lattice points.

4.4. Short chains

The next step in our program is to analyze the properties of representation (4.13) by calculating thermodynamic functions and correlation functions for small systems. Explicit expressions for the energy, specific heat and correlation functions are derived following the procedure given in chapter 2.

Recall that path-integral representation (4.13) gives exact results (albeit not in closed form) for any value of m for both the free-fermion ($v_1 = 0$) and the classical ($t = 0$) limit whereas for the checkerboard and real-space decomposition only the classical limit is recovered correctly. Thus in order to compare the different schemes we should look at the interacting system. We will choose $v_1 = 2t$ because this is a special point in the parameter space of (4.1). Indeed, for $v_1 = 2t$ the rotational symmetry in spin space of the corresponding spin-1/2 chain changes, model (4.1) maps onto the *isotropic* antiferromagnetic Heisenberg chain, or equivalently, there is a metal–insulator transition in the ground state of the fermion system if $\rho = \frac{1}{2}$. To study the convergence of the different approximation schemes we will confine the discussion to the "most difficult" case where the system is half filled, i.e. $\rho = \frac{1}{2}$. Indeed, from the relationship between (4.1) and (4.3) it follows that the ground state of the $\rho = \frac{1}{2}$ fermion system corresponds to the ground state of the antiferromagnetic spin chain and it is well-known that quantum effects play a very important role in determining the ground-state properties of this system.

For N spinless fermions on M sites the number of states s is given by the number of possibilities to

put the N particles on M sites, $s = \binom{M}{N}$, and consequently the Hamiltonian (4.1) can be represented by an $s \times s$ matrix. We will take $M \leq 8$ and $N \leq M/2$ (if $N > M/2$ we replace particles by holes) because diagonalization of matrices of dimension 70×70 does not require too much CPU-time.

We first look in more detail at numerical difficulties that may arise by employing path-integral form (4.13). From expression (4.13) it is immediately clear that each state in the sum contributes to Z_m^{PI}. A graphical representation of a typical configuration contributing to (4.13) is shown in fig. 4.1. There are no "paths" (straight lines connecting particles with the same label i and successive labels j) of zero weight but there are of course paths that give negative contributions to Z_m^{PI}. Although there is a large cancelation of positive and negative terms (see chapter 7 for more details) it is not difficult to arrange the computation such that these potential problems can be avoided.

In table 4.1 we show typical results for a system of 4 fermions on a ring of 8 sites. As for the 1D Ising model in a transverse field, we observe that here also the approximant to the free energy is larger than the approximant to the energy if m is small. Even though the temperature $T = 0.5$ is low, the convergence of the free energy and the energy is rather good. Taking $m = 16$ yields the exact energy within a relative error of less than 1%. A nice feature of the approximants F_m^{PI} and E_m^{PI} is that in all cases that we have studied, they converge monotonically towards the exact results. The monotonic increase of F_m^{PI} is in concert with the generalized Golden inequalities (see chapter 2). We do not know of any similar rigorous proof for the approximant E_m^{PI}. The approximant C_m^{PI} does not converge monotonically and in fact it converges slowly (for $m = 16$ the error is 8%). The rate of convergence of the approximant to the static structure factor is satisfying. The error for $m = 16$ is 1% which is roughly the same as for the energy. To illustrate the effect of lowering the temperature on the rate of convergence, table 4.2 shows the same four quantities as in table 4.1 for $T/t = 0.25$. The lower the temperature the larger is the value of m required to get a reasonable approximation, and in particular a very large m is necessary to get some accuracy on the specific heat. To get the same accuracy on the energy and structure factor as in the previous example (table 4.1), we would have to take $m = 32$. However for $m = 32$ the relative error on the specific heat is 2700%, even for $m = 256$ we still have an error of 85%. This example clearly demonstrates that changing the temperature and keeping β/m

Table 4.1
The path-integral approximation (4.13) to the free energy per site F_m^{PI}/M, energy per site E_m^{PI}/M, specific heat per site C_m^{PI}/M and structure factor $S_m^{\mathrm{PI}}(q = \pi)$ for a system of 4 spinless fermions on a ring of 8 sites as a function of the number of products m. The temperature $T/t = 0.50$ is fairly low as can be seen by comparing the exact ground-state energy per site $E_0 = -0.3685$ with the exact thermal energy E_∞^{PI}/M. The choice of the interaction $v_1 = 2t$ is such that spinless fermion model (4.1) is equivalent to the antiferromagnetic Heisenberg spin-1/2 model

m	F_m^{PI}/M	E_m^{PI}/M	C_m^{PI}/M	$S_m^{\mathrm{PI}}(q = \pi)$
1	−0.5232	−0.5765	0.2333	1.8501
2	−0.4603	−0.4796	0.3765	1.3668
3	−0.4379	−0.4272	0.2951	1.1468
4	−0.4283	−0.4017	0.2174	1.0518
8	−0.4176	−0.3714	0.1027	0.9493
16	−0.4147	−0.3626	0.0819	0.9218
32	−0.4140	−0.3603	0.0848	0.9149
∞	−0.4137	−0.3595	0.0892	0.9125

Table 4.2
Same as in table 4.1 except that here $T/t = 0.25$

m	F_m^{PI}/M	E_m^{PI}/M	C_m^{PI}/M	$S_m^{\mathrm{PI}}(q = \pi)$
1	−0.5603	−0.6031	0.0137	1.9973
2	−0.5010	−0.5808	0.3469	1.8655
3	−0.4609	−0.5304	0.6044	1.5969
4	−0.4375	−0.4867	0.5540	1.3861
8	−0.4050	−0.4102	0.1076	1.0698
16	−0.3942	−0.3802	−0.0996	0.9663
32	−0.3912	−0.3715	−0.0728	0.9385
∞	−0.3902	−0.3685	0.0028	0.9290

constant does not guarantee anything about the quality of the mth approximation as such. Only by a careful study of the convergence of the various quantities by means of exact short chain calculations can one get a feeling for the minimum value of m that is needed to reach a specified accuracy.

As the norms of the operators H_1 and H_2 (see (4.1b, c)) depend on the number of particles, (2.3) suggests that for a fixed density, the approximation might become less accurate if the number of sites increases. In table 4.3 we show the exact numerical results of a system of 2 particles and 4 sites for the same set of parameters as used for table 4.1. Except for small m we could not find evidence that for model (4.1), the approximation deteriorates by increasing the number of sites and keeping ρ fixed, a behavior for which we cannot offer a mathematical explanation.

In order to compare the path-integral approximation with the checkerboard and real-space approximation we only have to replace the algorithm for setting up the matrix $\mathscr{A}$ (see section 2.2) and repeat the calculations. In table 4.4 we give typical results for the checkerboard representation whereby we have chosen the same set of model parameters as in table 4.2. For fixed m we observe that for all quantities of interest the path-integral form is the most accurate approximation, the checker-board approximation the worst. Of course as $m \to \infty$ the particular choice of partitioning becomes irrelevant since all three approximations converge to the exact result. However in case we have to resort to the Monte Carlo method, taking the limit $m \to \infty$ will not be possible for practical reasons (see below) but since we know which representation will give us the best estimates we can choose the most efficient scheme.

We want to stress that short chain calculations of the kind we have done here, are as yet the only way to investigate quantitatively the properties of the various approximations. Also it is not easy to generalize conclusions drawn from the study of a specific case. For instance if we would use the similarity between the spinless fermion model and spin-1/2 model (4.3) to make the statement that the corresponding path-integral form of the spin-1/2 partition function would be the most accurate approximation, this would be wrong (De Raedt, Lagendijk and Fivez, 1982).

Table 4.3
Same as in table 4.1 except that here $N = 2$ and $M = 4$

m	F_m^{PI}/M	E_m^{PI}/M	C_m^{PI}/M	$S_m^{PI}(q=\pi)$
1	−0.5024	−0.4900	0.0863	0.9814
2	−0.4479	−0.4316	0.2599	0.8856
3	−0.4238	−0.3826	0.2359	0.8170
4	−0.4127	−0.3552	0.1596	0.7812
8	−0.4004	−0.3211	0.0319	0.7387
16	−0.3970	−0.3111	0.0095	0.7266
32	−0.3962	−0.3085	0.0138	0.7235
∞	−0.3959	−0.3076	0.0187	0.7224

Table 4.4
The Barma and Shastry checkerboard approximation (4.19) to the free energy per site F_m^{CBD}/M, energy per site E_m^{CBD}/M, specific heat per site C_m^{CBD}/M and structure factor $S_m^{CBD}(q=\pi)$ for a system of 4 spinless fermions on a ring of 8 sites as a function of the number of products m. The temperature $T/t = 0.25$. Comparison with table 4.2 shows that the path-integral approximation is superior

m	F_m^{CBD}/M	E_m^{CBD}/M	C_m^{CBD}/M	$S_m^{CBD}(q=\pi)$
1	−0.6851	−0.7497	0.0108	1.9983
2	−0.6020	−0.7347	0.2670	1.9100
3	−0.5348	−0.6822	0.8038	1.6812
4	−0.4889	−0.6129	1.0065	1.4574
8	−0.4203	−0.4549	0.1708	1.0860
16	−0.3981	−0.3919	−0.2001	0.9693
32	−0.3922	−0.3744	−0.1229	0.9391
∞	−0.3902	−0.3685	0.0028	0.9290

4.5. Monte Carlo simulations

If the number of states of spinless fermion model (4.1) is larger than let us say thousand, it becomes very hard to calculate the model properties by means of numerically exact methods. The only useful technique that can handle problems of this size is the Monte Carlo method discussed in chapter 2. In this section we address the additional, more specific, difficulties that are encountered if the Monte Carlo scheme is utilized to simulate systems described by the path-integral or checkerboard representation.

4.5.1. Path-integral representation

Let us first look at the path-integral form (4.13). As there are negative contributions to Z_m^{PI}, we cannot simply use (4.13) to define a transition probability. However, for the spinless fermion model this fundamental "minus-sign" problem can be circumvented by employing the intimate relationship with other models (De Raedt and Lagendijk, 1981).

It is obvious that the artificial system described by the approximant

$$\tilde{Z}_m^{\mathrm{PI}} = \sum_{\{i_{1,j}<\cdots<i_{N,j}\}} \sum_{\{P_j\}} \prod_{j=1}^{m} \prod_{\mu,\nu=1}^{N} I\Big(\frac{2\beta t}{m}, i_{\mu,j} - i_{P_j\mu,j+1}\Big) \exp\Big(-\frac{\beta v_1}{m} \delta_{|i_{\mu,j}-i_{\nu,j}| \bmod M,1}\Big), \tag{4.26}$$

can be simulated by means of the Monte Carlo method. As the only difference between (4.26) and (4.13) is that the sign function has been removed, all terms in (4.26) are positive and consequently it is straightforward to define the positive transition probabilities (i.e. $g(\mathcal{S})$ in section 2.3) that we need to construct a Markov chain of states (here "state" means the set of all particle positions on the 2D lattice plus permutations). We denote expectation values of operators taken in this artificial ensemble by $\langle\langle \cdot \rangle\rangle$. It is trivial to show that all (except the free energy) physical quantities of the spinless fermion system can be written as the ratio of these double bracket expectation values. For example the approximant to the energy reads

$$E_m^{\mathrm{PI}} = \langle\langle \mathrm{sign}(P_1 \cdots P_m) e_m \rangle\rangle / \langle\langle \mathrm{sign}(P_1 \cdots P_m) \rangle\rangle, \tag{4.27a}$$

where

$$e_m = \frac{t}{m} \sum_{\mu=1}^{N} \sum_{j=1}^{m} \frac{I(2t\beta/m, i_{\mu,j}-1) + I(2t\beta/m, i_{\mu,j}+1)}{I(2t\beta/m, i_{\mu,j})} - \frac{v_1}{m} \sum_{\mu,\nu=1}^{N} \sum_{j=1}^{m} \delta_{|i_{\mu,j}-i_{\nu,j}| \bmod M,1}. \tag{4.27b}$$

Remains the question whether statistical errors on expectation values such as $\langle\langle \mathrm{sign}(P_1 \cdots P_m) \rangle\rangle$ will be small. In general the answer to this question is certainly no but for the problem at hand the answer is affirmative because our seemingly heuristic approach to the minus-sign problem is founded on a sound theoretical basis.

First we note that except for the restriction on the sum over all possible positions of the particles (4.27) is identical to the mth approximation

$$\hat{Z}_m^{\mathrm{PI}} = \sum_{\{i_{\mu,j}\}} \sum_{\{P_j\}} \prod_{j=1}^{m} \prod_{\mu,\nu=1}^{N} I\Big\{\frac{2\beta t}{m}, i_{\mu,j} - i_{P_j\mu,j+1}\Big) \exp\Big(-\frac{\beta v_1}{m} \delta_{|i_{\mu,j}-i_{\nu,j}| \bmod M,1}\Big), \tag{4.28}$$

to the boson partition function. Let us denote a state of N bosons on M sites by $b_{i_1}^+ \cdots b_{i_N}^+ |0\rangle$, $l_\mu = 1, \ldots, M$ and write $\mathcal{P}$ for the projector on the subspace $\{b_{i_1}^+ \cdots b_{i_N}^+ |0\rangle;\ l_\mu = 1, \ldots, M, \mu \neq \nu \Rightarrow l_\mu \neq l_\nu\}$. Then we have

$$\hat{Z}_m^{\mathrm{PI}} = \mathrm{Tr}[\exp(-\beta \hat{H}_1/m) \exp(-\beta \hat{H}_2/m)]^m , \tag{4.29a}$$

and

$$\tilde{Z}_m^{\mathrm{PI}} = \mathrm{Tr}[\mathcal{P} \exp(-\beta \hat{H}_1/m) \mathcal{P} \exp(-\beta \hat{H}_2/m) \mathcal{P}]^m , \tag{4.29b}$$

where $\hat{H}_1 = -t \sum_{i=1}^{M} (b_i^+ b_{i+1} + b_{i+1}^+ b_i)$ and $\hat{H}_2 = v_1 \sum_{i=1}^{M} b_i^+ b_i b_{i+1}^+ b_{i+1}$ stand for the boson kinetic energy and potential energy respectively. Introducing the projection operator $\mathcal{P}$ is the same as saying that there is a "hard core" interaction that prevents two particles from occupying the same site. Note that $\mathcal{P}^2 = \mathcal{P}$ so that the number of $\mathcal{P}$'s in (4.30b) is $2m$ rather than $3m$. Another way to take this hard core interaction into account is to define a projected boson Hamiltonian by $\bar{H} = \mathcal{P}(\hat{H}_1 + \hat{H}_2)\mathcal{P}$. The corresponding Trotter approximation reads

$$\bar{Z}_m^{\mathrm{PI}} = \mathrm{Tr}[\mathcal{P} \exp(-\beta \mathcal{P} \hat{H}_1 \mathcal{P}/m) \mathcal{P} \exp(-\beta \mathcal{P} \hat{H}_2 \mathcal{P}/m) \mathcal{P}]^m , \tag{4.30}$$

and it is clear that in general $\bar{Z}_m^{\mathrm{PI}} \neq \hat{Z}_m^{\mathrm{PI}} \neq \tilde{Z}_m^{\mathrm{PI}} \neq \bar{Z}_m^{\mathrm{PI}}$ but we do have $\lim_{m\to\infty} \hat{Z}_m^{\mathrm{PI}} \neq \lim_{m\to\infty} \tilde{Z}_m^{\mathrm{PI}} = \lim_{m\to\infty} \bar{Z}_m^{\mathrm{PI}}$. In the limit $m \to \infty$ the artificial model $\tilde{Z}_m^{\mathrm{PI}}$ that will be used in the Monte Carlo simulation and the projected model Z_m^{PI} coincide. Letting the boson operators act on the subspace defined by the projection $\mathcal{P}$ is the same as working with spin-1/2 Pauli matrices if we replace $b_l^+ \to \sigma_l^+$, $b_l \to \sigma_l^-$ and $b_l^+ b_l \to (1 + \sigma_l^z)/2$. Then it follows immediately that H is exactly the same as Hamiltonian (4.3). Thus we conclude that the artificial model that we use in the Monte Carlo simulation is closely related (and becomes identical in the $m \to \infty$ limit) to the spin-1/2 model (4.3). Therefore it is not really a surprise that it makes sense to simulate artificial model (4.23) and "measure" fermion quantities because the system that we actually simulate is closely related to the spinless fermion model and behaves in many respects as the model that we want to investigate. Clearly, if this approach works, it only works because there is a one-to-one correspondence between 1D spinless fermion model (4.1), hard-core boson Hamiltonian $\hat{H}_1 + \hat{H}_2$ and 1D spin-1/2 model (4.3). For 2D or 3D this relationship is lost and one should not expect that the minus-sign trick will work there. As a matter of fact, for the two-dimensional spinless fermion model it does not.

Having found a way to deal with the minus signs, we now briefly summarize how we simulate the system defined by (4.27). Compared with a simulation of a genuine classical model, the only complication is that in addition to the particle positions we also have to consider permutations as variables. The actual algorithm that we use is simple. The $N \times m$ particles move on a two-dimensional $M \times m$ lattice. We pick out a new trial state by choosing the particle index μ and the Trotter index j at random and generate a trial position for this particle. In choosing this trial position we have to take into account that the number of particles on each replica of the original quantum chain is conserved or, in other words, a particle can only jump in the real-space direction. Now there are two distinct possibilities. If the new trial position is not occupied we calculate the transition probability for this jump and accept the jump with this probability. If there was already another particle on the trial position we permute the two particles, calculate the change in the contribution to the partition function and accept the new permutation (of which the index is j) with the corresponding probability. More sophisticated extensions

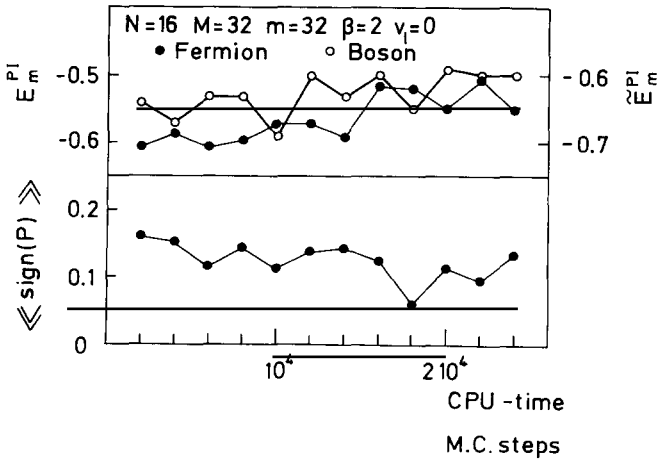

Fig. 4.4. Average sign of the permutation operator $P = P_1 \cdots P_m$, energy per site of the 1D fermion model E_m^{PI} and energy per site of the "boson" model $\tilde{E}_m^{\mathrm{PI}}$ as a function of the number of Monte Carlo steps.

of this algorithm have been tried out but we have found the most simple version adequate for our purposes.

In Monte Carlo studies it is always necessary to look at the dependence of the quantities sampled during the simulation on the simulation time (i.e. the number of Monte Carlo steps or CPU time). In fig. 4.4 we display some typical results of a simulation of 16 particles on 32 sites. We have plotted $\langle\langle \mathrm{sign}(P_1 \cdots P_m)\rangle\rangle$, the fermion energy E_m^{PI} and the "boson" energy $\tilde{E}_m^{\mathrm{PI}}$ as a function of the number of Monte Carlo steps per "particle" (16×32 in this case). Each point represents the average of 2000 samples, the number of Monte Carlo steps between taking samples being mN. We conclude that the Monte Carlo procedure is quite stable.

4.5.2. Checkerboard representation

We first restrict the discussion to the case of free boundary conditions (in the real dimension, not in the Trotter direction). Then there are no negative contributions to Z_m^{CBD} (or Z_m^{RSD}, remember CBD stands for checkerboard decomposition and RSD for real-space decomposition) if the quantum system is one-dimensional and we can concentrate on the specific problem of this approach, the presence of "forbidden" configurations. From (4.20, 21) it follows directly that some configurations $S_{i,j}$, $\bar{S}_{i,j}$ will have zero weight because one or more matrix elements in the product (4.19) are zero. Note that the presence of zero's in matrix (4.20) results from the local conservation of the number of particles, i.e. each two-site Hamiltonian conserves the number of spinless fermions.

As more than one half of the elements of matrix (4.20) are zero one might expect that the fraction of forbidden configurations is large. More insight is gained by using the formulation where we summed out the intermediate variables $\bar{S}_{i,j}$. From (4.21) it follows that the product of T's in (4.22) vanishes for configurations $\{S_{i,j}, \sigma_j\}$ such that there is at least one non-diagonal element of T_{-1}. As a matter of fact it turns out that it is very cumbersome to find configurations that have non-zero weight, especially if m becomes large. Exact summations for short chains (De Raedt and Lagendijk, 1982a, 1982b; De Raedt, Lagendijk and Fivez, 1982) demonstrate that the fraction of non-zero contributions to Z_m^{CBD} vanishes as $m \to \infty$. Obviously this is a serious problem if one intends to use the Monte Carlo method because it would be extremely inefficient to generate all these forbidden states during the simulation.

Since this technical difficulty is due to conservation laws, the only way to avoid generation of useless configurations is to build these conservation laws into the algorithm that constructs trial states from a previous, allowed state. This excludes the use of one-particle jump schemes as they do not keep the number of particles in each two-site block (each shaded area in figs. 4.2, 4.3) constant. Thus the minimal local change that can be made is to change the position of two particles at a time.

An ingeniously simple scheme (Hirsch et al., 1982) that accomplishes this is to move two particles from one vertical edge of an unshaded box to the other, as is shown in fig. 4.5. The open circles in fig. 4.5 indicate the new positions of the two particles after the move. From fig. 4.5 it is also clear that it is not always possible to make this kind of move. Assuming the move described above has been accepted, it is no longer possible to move any pair of particles which is currently on spatial site 3 to the left. In sweeping over the lattice one has to verify whether it is allowed to move the two particles from one edge to the other. One would therefore expect that the local two-particle jump scheme becomes inefficient if the density $\rho \simeq \frac{1}{2}$ (if $\rho > \frac{1}{2}$ use particle–hole symmetry) and if m is not too small and one expects that there will be strong correlations between successive states. As far as we know no detailed study of these potential deficiencies of the checkerboard approach to the 1D fermion problem has been reported. However the checkerboard approximation has also been used for the closely related spin-1/2 chain. For the spin-1/2 model the Monte Carlo scheme is more complicated than for the spinless fermion model because one also has to sum over all possible magnetizations (number of particles in fermion language) and to do this requires an additional non-local spin-flip procedure (see also chapter 5). One observes that if $m > 8$ it becomes very time consuming to find a pair of particles that can be moved (Cullen and Landau, 1983). Since it is not easy to find two particles that can be moved this algorithm also induces correlations between successive states and indeed for the spin-1/2 model it was found that runs with different random sequences gave completely different results for the susceptibility (Cullen and Landau, 1983). We want to stress that these drawbacks of the checkerboard approach is the consequence of trying to build into the algorithm, the conservation of the local number of particles and is not in any sense, the result of a particular choice of physical parameters such as the temperature or nearest-neighbor interaction strength.

In Monte Carlo studies of lattice model such as (4.1) one normally likes to impose periodic boundary conditions in the real-space direction (recall that because of the trace operation we may not choose anything else than periodic boundary conditions in the Trotter direction). One then hopes that the properties of the system will resemble more closely those of the infinite system. From (4.19) it follows

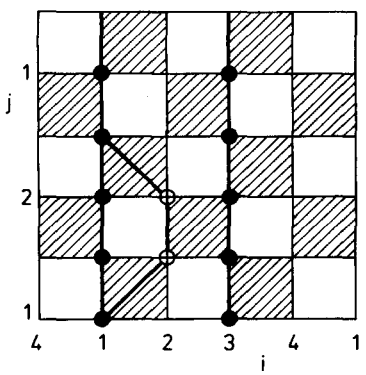

Fig. 4.5. Example of the elementary two-particle jump procedure for the checkerboard lattice. A trial state is generated by moving two particles from one vertical edge of an unshaded box to the other.

immediately that if we want to simulate the checkerboard lattice with periodic boundary conditions, there will be negative contributions to Z_m^{CBD} if the number of particles is even. Obviously for a 1D fermion model this problem can be circumvented by confining oneself to simulations of systems in which the number of particles is odd.

A prerequisite for any Monte Carlo scheme is that the procedure for generating trial states can, at least in principle, bring the system from one state to another (see section 2.3). It is fairly easy to see that the local two-particle jump scheme proposed by Hirsch et al. (Hirsch et al., 1981, 1982) can generate all states of the checkerboard lattice with *free* boundary conditions. However, in the case of *periodic* boundary conditions this scheme has to be extended in order to guarantee that any state of the system can be reached after a possibly large, but finite, number of moves. The reason for this is that by imposing periodic boundary conditions, one allows for a class of configurations for which the order in which particles are created from the vacuum state, depends on the chain (Trotter) index j and on the number of particles that cross the real-space boundaries. If we follow the path (the trajectory as a function of the Trotter index j) of the lth particle on chain $j = 1$ for such a configuration, we find that we end up at the $(l+n)$th particle at chain $j = m+1$ $(=1)$. The number n is positive if the particle moves to the right when j changes from 1 to $m+1$ and is negative otherwise. If $n = \pm 1$, the net number of particles crossing the boundaries is one. This necessarily implies that if we go from chain $j = 1$ to chain $j = m$ and back to $j = m+1$ $(=1)$ we have to resequence the labels of the particles and if the number of particles is even, such a configuration will have a negative weight. In general if N is even only configurations with n even have a positive weight. More important from the point of view of Monte Carlo simulation is that it is impossible to change n if one only has the two-particle jump procedure available. For instance going from an $n = 0$ state to an $n = \pm 1$ state requires that one must allow for moving N particles in one step. In principle such multi-particle moves satisfy the conditions for a correct Monte Carlo scheme but in practice it is not easy to devise an algorithm that gives sufficient coverage of configuration space and yields a rate for accepting new configurations which is not too low.

To avoid the problem of having to generate N-particle moves and the difficulties caused by negative contributions at the same time Hirsch et al. (Hirsch et al., 1981, 1982) have chosen to simulate the checkerboard lattice with periodic boundary conditions by making use of two-particle moves only. By starting from an $n = 0$ configuration, this scheme necessarily keeps $n = 0$ and consequently it is impossible to reach a state with negative weight or $n \neq 0$. It is obvious that this Monte Carlo method cannot generate all the states of the system. We will call this approximation the "modified checkerboard decomposition" (MCBD). On the other hand it is clear that the minus-sign problem and the non-ergodicity of the Markov chain resulting from the application of the two-particle jump procedure are solely due to the requirement to simulate a system with periodic boundary conditions. One expects that boundary effects are unimportant unless a typical correlation length exceeds the size of the system being simulated. For 1D model (4.1) we know that there only can be long range order if $T = 0$, $\rho = \frac{1}{2}$ and $v_1 > 2t$. With Monte Carlo techniques based upon the Trotter formula one can only perform simulations at $T \neq 0$ and consequently there can at most be "incipient" long range order in the 1D system. In practice the simulation method itself can be used to test whether or not the system is large enough. One has to repeat the simulation for chains of different length. Thereby one should take into account that both m and M have to be chosen judiciously. From tables 4.1–4.4 we see that $S_m^{\mathrm{PI}}(q = \pi)$ and $S_m^{\mathrm{CBD}}(q = \pi)$ schematically overestimate the exact value of $S(q = \pi)$. As $S(q = \pi)$ is a measure for the correlation between particles, choosing m too small will give the impression that the correlation length relevant for the simulation itself, is larger than it should be.

Monte Carlo simulation data for spinless fermion model (4.1) obtained from the MCBD can be found in (Hirsch et al., 1981, 1982). Of particular interest are the MCBD simulation data for the structure factor (see fig. 9c of Hirsch et al., 1982). The simulations have been done such that $m = M$ and $\beta/m = 0.25$ (Hirsch, 1983c) and one observes that $S(q = \pi)$ grows logarithmically with the β/M. Hirsch et al. (Hirsch et al., 1982) interpret this logarithmic dependence as evidence that their simulation data is in qualitative agreement with the prediction based upon a ground-state calculation of the continuum version of the spinless fermion model (Luther and Peschel, 1975). On the other hand, simulation data for the PI do not show any detectable size dependence if $M \geq 8$ (De Raedt and Lagendijk, 1981, 1982a). The difference in interpretation stems from the fact that Hirsch et al. are comparing data for systems of different size and different temperatures (but $\beta/M = \beta/m = 0.25$) whereas we kept the temperature fixed and changed M and increased m until the improvement in accuracy of the approximants became smaller than the statistical noise in the data.

4.6. Applications

Up to now only the simple spinless 1D fermion model has been treated in detail. This model served the purpose of illustrating the method. In this section we will discuss some more complicated models.

In the first place we will extend the 1D spinless fermion model to include more-distant-neighbor interactions. We generalize Hamiltonian (4.1) by replacing H_2 by

$$H_2 = \frac{1}{2} \sum_{i,j=1}^{M} v_{|i-j|} n_i n_j . \tag{4.31}$$

The inclusion of competing interactions might have dramatic consequences for the ground state of the classical model defined by (4.31) (Hubbard, 1978). It has been shown that under certain restrictions on the interactions v_l the ground state is periodic (Hubbard, 1978; Pokrovsky and Uimin, 1978). The period as well as the arrangement of the particles within each period depend on the particle density ρ. These (classical) ground-state configurations may be regarded as generalized one-dimensional Wigner lattices. According to Hubbard this classical model might offer a possible explanation for the optical spectra of tetracyanoquino (TCNQ) salts and the satellites observed in X-ray and neutron diffraction experiments (Pouget et al., 1976, 1979; Kagoshima et al., 1976). As pointed out by Hubbard, it was not clear whether his conclusions would remain valid if the finite bandwidth would have been taken into account.

Extension (4.31) is trivially included in the path-integral method (De Raedt and Lagendijk, 1983a). Detailed information about the arrangement of the particles on the chain can be extracted from the static structure factor $S(q)$ and the static (density–density) susceptibility $\chi(q)$. The effect of the next-nearest-neighbor interactions on $S(q)$ (see (4.4b)) for $\rho = \frac{1}{2}$ is shown in fig. 4.6. The inclusion of v_2 does induce a change in $S(q)$. The single-peak structure changes into a broad double-peak structure indicating that the nature of the most probable particle arrangement changes when v_2 increases (while keeping other parameters constant). The very strong and narrow peaks of the classical model do not resemble those of the quantum mechanical system.

In fig. 4.7 we show the effect on turning on v_3. The set of parameters ($\rho = 13/32$, $v_1 = 7$, $v_2 = 3$, $v_3 = 1$) were chosen so as to correspond to TTF-TCNQ (Hubbard, 1978). It is interesting that the introduction of v_3 makes the peaks in $S(q)$ less prominent and less like the classical analogon. It is clear that the hopping term has a dramatic effect. All the exotic features of the classical model like the occurrence of higher harmonics in the classical $S(q)$ (Kondo and Yamaji, 1984) are wiped out completely by the quantum fluctuations.

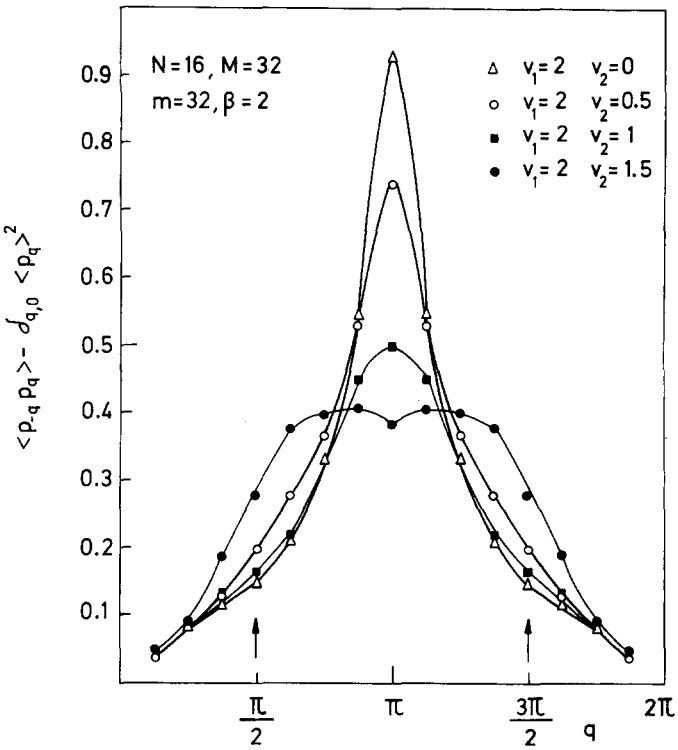

Fig. 4.6. Influence of the next-nearest-neighbor interaction on $S(q)$. The arrows give the peak positions of $S(q)$ in the classical limit ($t = 0$) for the case $2v_2 > v_1$. Statistical errors are less than 5%. Solid lines through the data points are guides to the eyes only.

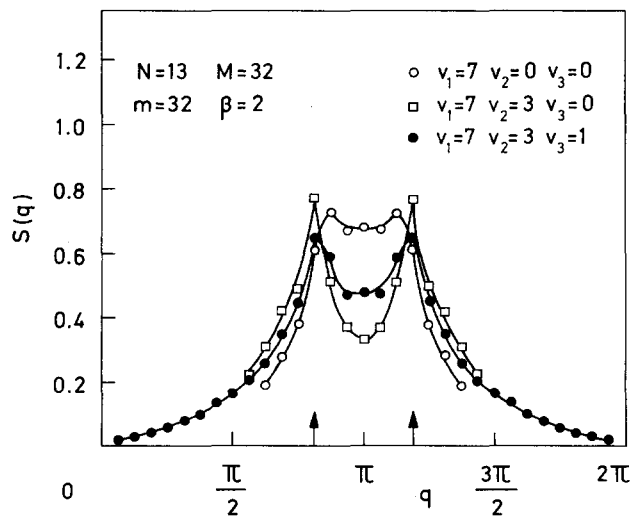

Fig. 4.7. The static structure factor $S(q)$ of a one-dimensional system of spinless fermions for several values of the near-neighbor interactions v_l. Solid lines through the data points are guides to the eyes only.

Introducing more-distant interactions within the RSD and CBD partitionings is more difficult than in the path-integral approach because the former partitions use the local nature of the model and the more-distant interactions destroy the extreme locality. It is of course possible that a combination of the path-integral method and a real-space type of partitioning might be useful but so far no results obtained from such an approach have been published.

A very important extension from the point of view of applications is the inclusion of spin. A possible way to do this is to follow the approach of Hirsch et al. (Hirsch et al., 1981, 1982) who suggest to treat the model in an ensemble where both the number of up and down spins are held constant. A number of important questions need to be settled before one can use this method. For instance one must be absolutely sure that the true ground state of the system belongs to the subspace spanned by the states of the restricted ensemble. Furthermore experimental conditions on real systems might need to be explained using a different ensemble.

The most important 1D fermion model is certainly the extended Hubbard model

$$H = -t \sum_{i=1}^{M} \sum_{\sigma=\uparrow,\downarrow} (c^+_{i,\sigma} c_{i+1,\sigma} + c^+_{i+1,\sigma} c_{i,\sigma}) + U \sum_{i=1}^{M} \sum_{\sigma=\uparrow,\downarrow} c^+_{i,\sigma} c_{i,\sigma} c^+_{i,-\sigma} c_{i,-\sigma} + V \sum_{i,j=1}^{M} \sum_{\sigma,\sigma'=\uparrow,\downarrow} c^+_{i,\sigma} c_{i,\sigma} c^+_{j,\sigma'} c_{j,\sigma'} . \tag{4.32}$$

This model plays a central role in the theory of the one-dimensional electron gas (Emery, 1979). Adopting the Barma-and-Shastry CBD break-up with periodic boundary conditions and neglecting all negative contributions, this model has been studied extensively by means of the Monte Carlo simulation technique (Hirsch et al., 1981, 1982; Hirsch and Scalapino, 1983a, 1983b). The method has been tested on the phase diagram of the quarter-filled ($\rho = \frac{1}{2}$) band.

Of the various correlation functions that have been sampled, we mention staggered spin-density, charge-density and singlet-superconducting correlation functions (Hirsch et al., 1982). This study of the $\rho = \frac{1}{2}$ case was extended further by Hirsch and Scalapino (Hirsch and Scalapino, 1983a, 1983b). In this work susceptibilities rather than correlation functions were measured. One conclusion drawn for these data is that the Coulomb interactions suppress the $2k_F$ peak in the charge-density susceptibility and that the $2k_F$ spin-density susceptibility is greatly enhanced by a finite U, and even more if $V \neq 0$. To induce a maximum at $4k_F$ in the charge-density response function requires $V > 0$. Model (4.32) was studied further by Hirsch and Scalapino (Hirsch and Scalapino, 1983b) for the more realistic and also more complicated case $\rho = 0.6$. Again the presence of both a finite U and V has an important influence on the relative strength of the $2k_F$ and $4k_F$ susceptibilities. The authors also discuss the relevance of their work with respect to experiments on TCNQ compounds. As a detailed discussion of this vast amount of work requires a profound knowledge of the physics of quasi-1D compounds, it is far beyond the scope of this review and we refer the reader to the original papers for more details.

5. Spin-1/2 models

In this chapter we introduce path-summation representations for partition functions of spin-1/2 lattice models by explicitly using the local and lattice structure of the theory. The model Hamiltonian is given by

$$H = -\sum_{\langle ij \rangle} H_{i,j} , \tag{5.1a}$$

$$H_{i,j} = J_1(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y) + J_3 \sigma_i^z \sigma_j^z + \frac{\boldsymbol{h} \cdot (\boldsymbol{\sigma}_i + \boldsymbol{\sigma}_j)}{2}, \tag{5.1b}$$

where the sum over i and j runs over all nearest-neighbor bonds of the d-dimensional lattice. For several special choices for the exchange interactions J_1 and J_3 and the external field $\boldsymbol{h}$ the model exhibits global rotational symmetries in spin space. For $J_1 = J$, $J_3 = 0$ and $\boldsymbol{h} = h\boldsymbol{e}_z$ we recover the XY-model in a longitudinal external field. This XY-Hamiltonian remains invariant for rotations (in spin space) around the z-axis. The Heisenberg model is obtained by putting $J_1 = J_3 = J$ and has full rotational symmetry in spin space if $\boldsymbol{h} = 0$. The Ising model in a transverse field, treated in chapter 3, corresponds to the case $J_1 = 0$ and $\boldsymbol{h} = h\boldsymbol{e}_x$ and is only invariant for rotations by π around the x-axis. The general model (5.1) is itself a special case of the X-Y-Z Hamiltonian in which also the exchange interaction in x- and y-direction differ.

The symmetry properties of the full Hamiltonian have to be taken into consideration when one chooses a particular partitioning. For model (5.1), by writing the Hamiltonian as a sum over two-spin blocks a partitioning is achieved without changing the symmetry in spin space. Of course it is interesting to know what will happen if we break the original symmetry by choosing another decomposition. In the one-dimensional case we can study the consequences of such a choice in detail. In the Heisenberg limit $J_1 = J_3$ and for zero field $\boldsymbol{h} = 0$, partitioning the Hamiltonian into XY-model and Ising model obviously breaks the full rotational symmetry (which is of course restored when $m \to \infty$). Remark that this type of decomposition is the same as the one used for the Ising model in a transverse field (see chapter 3). All the eigenvalues and eigenvectors of the one-dimensional spin-1/2 XY Hamiltonian are known (Lieb, Schultz and Mattis, 1961; Katsura, 1962; Lieb and Mattis, 1966) and are very similar to those of a one-dimensional free fermion model. Thus it is not surprizing that one may derive a path-summation representation for the one-dimensional model (5.1) which is almost identical to (4.13), the only differences being that one has to sum over all possible magnetizations (particles in the fermion problem) and replace (4.11) by a slightly different formula (De Raedt, Lagendijk and Fivez, 1982). It is not possible to extend this analogy between fermion and spin path-sum representations to higher dimensions simply because the eigenvalues and eigenvectors of the two- and three-dimensional spin-1/2 XY model are not known.

5.1. One-dimensional model

To proceed further it is necessary to specify the model in more detail because otherwise we would get lost in the complexity of the notation. Therefore we will first confine ourselves to the one-dimensional case, choose $\boldsymbol{h} = h\boldsymbol{e}_z$ and work in the representation that diagonalizes σ_i^z ($\sigma_i^z|S_i\rangle = S_i|S_i\rangle$). If we consider the same orderings of the two-site blocks as in the case of the one-dimensional fermion system (see 4.2, 3) it is not difficult to see that (4.18) and (4.24) can be taken over without modification and in the case of the checkerboard decomposition one has to replace (4.19) by

$$\begin{aligned} Z_m^{\mathrm{CBD}} = \sum_{\{S_{i,j}\}} \sum_{\{\bar{S}_{i,j}\}} \prod_{j=1}^{m} & T(S_{1,j}, S_{2,j}; \bar{S}_{1,j}, \bar{S}_{2,j}) \cdots T(S_{M-1,j}, S_{M,j}; \bar{S}_{M-1,j}, \bar{S}_{M,j}) \\ & \times T(\bar{S}_{2,j}, \bar{S}_{3,j}; S_{2,j+1}, S_{3,j+1}) \cdots T(\bar{S}_{M-2,j}, \bar{S}_{M-1,j}; S_{M-2,j+1}, S_{M-1,j+1}) \\ & \times T(\bar{S}_{M,j}, \bar{S}_{1,j}; S_{M,j+1}, S_{1,j+1}), \end{aligned} \tag{5.2}$$

and a similar result is obtained for the case of the real-space partitioning. It is easy to find the explicit

expression for the T-matrices appearing in (5.2). We have

$$T(S_i, S_j; \bar{S}_i, \bar{S}_j) = \langle S_i, S_j | \exp\left(-\frac{\beta H_{i,j}}{m}\right) | \bar{S}_i, \bar{S}_j \rangle$$

$$= \exp(-\beta J_3/m) \begin{pmatrix} \exp\left(\frac{\beta(2J_3-h)}{m}\right) & 0 & 0 & 0 \\ 0 & \cosh\frac{2\beta J_1}{m} & \sinh\frac{2\beta J_1}{m} & 0 \\ 0 & \sinh\frac{2\beta J_1}{m} & \cosh\frac{2\beta J_i}{m} & 0 \\ 0 & 0 & 0 & \exp\left(\frac{\beta(2J_3+h)}{m}\right) \end{pmatrix} \begin{matrix} |-1\ -1\rangle \\ |\ 1\ -1\rangle \\ |-1\ \ 1\rangle \\ |\ 1\ \ 1\rangle \end{matrix}; \quad . \tag{5.3}$$

A peculiar feature of this type of decompositions is that the approximants for the partition function obtained from the checkerboard and real-space break-up are identical (De Raedt, Lagendijk and Fivez, 1982).

As for the fermion problem the sum over the $\bar{S}_{i,j}$ variables can be carried out analytically and we obtain

$$Z_m^{\mathrm{CBD}} = \sum_{\{S_{i,j}\}} \sum_{\{\sigma_j\}} \prod_{j=1}^{m} T_{\sigma_j\phi_{1,j}}(S_{1,j}, \sigma_j\phi_{1,j}S_{2,j}) \times T_{\sigma_j\phi_{2,j}}(\sigma_j\phi_{2,j}S_{2,j}, S_{3,j+1}) \cdots T_{\sigma_j\phi_{M-1,j}}(S_{M-1,j}, \sigma_j\phi_{M-1,j}S_{M,j}) T_{\sigma_j\phi_{M,j}}(\sigma_j\phi_{M,j}S_{M,j}, S_{1,j+1}), \tag{5.4a}$$

where

$$T_1(S, \bar{S}) \equiv \begin{pmatrix} \exp(\beta(J_3-h)/m) & \exp(-\beta J_3/m)\cosh(2\beta J_1/m) \\ \exp(-\beta J_3/m)\cosh(2\beta J_1/m) & \exp(\beta(J_3+h)/m) \end{pmatrix}; \quad \begin{matrix} |-1\rangle \\ |\ 1\rangle \end{matrix}, \tag{5.4b}$$

$$T_{-1}(S, \bar{S}) \equiv \delta_{S,\bar{S}} \exp(-\beta J_3/m)\sinh(2\beta J_1/m). \tag{5.4c}$$

The similarity between (5.2–4) and (4.19–22) is more than just a coincidence. If we neglect the effect of the periodic boundary condition in (4.19) (remove the minus-sign factor in (4.19)) and calculate the grand-canonical partition function of the fermion system instead of the canonical one (multiply (4.19) by $\exp(\beta\mu N)$ and sum over $N = 0, \ldots, M$) we see that the approximation to the grand-canonical partition function of the fermion system is identical (up to a trivial factor $\exp[\beta M(J_3 - h)]$) to the partition function of the one-dimensional spin-1/2 system if we make the identification $J_1 = t/2$, $J_3 = -v_1/4$ and $h = \mu/2 + v_1$. Of course the obtained relationship is not new since it is well-known that the Jordan–Wigner transformation maps the spin-1/2 chain onto the spinless fermion chain but it is remarkable that by using the Trotter formula one can prove this equivalence by only calculating the matrix representations of the exponents of the two-site Hamiltonians thus without invoking the non-linear and non-local Jordan–Wigner transformation.

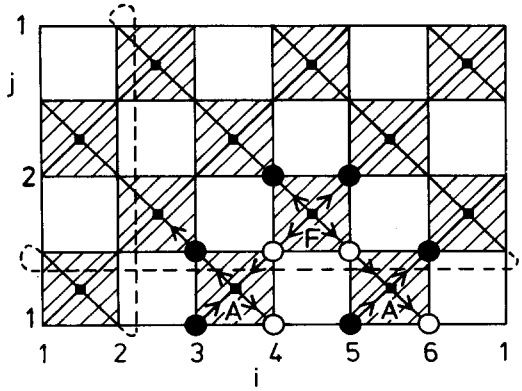

Fig. 5.1. Graphical construction for mapping the checkerboard representation of the spin-1/2 chain onto an 8-vertex model for the case of a ring of $M = 6$ spins and $m = 2$. Some of the Ising spins are shown explicitly: spin-up (down) corresponds to a full (open) circle. The four Ising spins on the corners of each shaded square interact with each other. For the Heisenberg and XY models a vertex configuration denoted A are allowed, the one labeled F is forbidden because the corresponding weight is zero. The boundary conditions of the 8-vertex model are indicated by dashed lines.

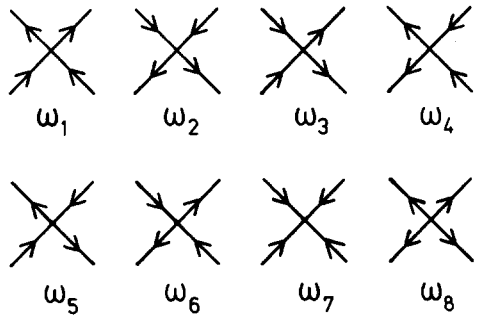

Fig. 5.2. The eight vertex configurations and their corresponding weights.

As for the Ising model in a transverse field it is possible to interpret the representation (5.2) in terms of a well-known two-dimensional lattice model, an 8-vertex model. This equivalence follows directly from the graphical construction invented by Barma and Shastry (Barma and Shastry, 1978) and is reproduced in fig. 5.1. As for the case of one-dimensional lattice fermions, we can interpret the checkerboard representation of the spin-1/2 problem in terms of a two-dimensional effective lattice by putting Ising spins on each corner of the shaded squares. The four Ising spins on each square interact with each other through the four-spin interaction matrix given by (5.3). Drawing the diagonals of all shaded squares we can turn the Ising spin lattice into a vertex model by assigning vertex configurations to spin configurations such that the arrow on the line through the site of the Ising spin run upwards (downwards) if the Ising spin is up (down). It is easy to check that the vertex model is indeed an 8-vertex model with weights (see fig. 5.2) given by

$$\omega_1 = \exp(\beta(J_3 + h)/m)\,, \tag{5.5a}$$

$$\omega_2 = \exp(\beta(J_3 - h)/m)\,, \tag{5.5b}$$

$$\omega_3 = \omega_4 = \exp(-\beta J_3/m)\cosh(2\beta J_1/m)\,, \tag{5.5c}$$

$$\omega_5 = \omega_6 = \exp(-\beta J_3/m)\sinh(2\beta J_1/m)\,, \tag{5.5d}$$

$$\omega_7 = \omega_8 = 0\,. \tag{5.5e}$$

In (5.1) the exchange interaction for x- and y-components is the same (J_1) and as a consequence of this symmetry two out of eight vertices have zero weight. However it is not difficult to generalize to the most general case (Barma and Shastry, 1978). Thus in fact the checkerboard approximation for the model (5.1) is equivalent to a 6-vertex model. From fig. 5.1 it is clear that the periodic boundary condition in the Trotter direction, which is a direct consequence of the trace operation and can therefore not be chosen at will, implies a rather odd boundary condition on the vertices.

It is well-known that there exists a remarkable similarity of the rigorously solvable two-dimensional classical lattice problems and one-dimensional system (5.1). In general this connection has been established by demonstrating that the transfer matrix of the 8-vertex model commutes with the X-Y-Z Hamiltonian (McCoy and Wu, 1968; Sutherland, 1970; Baxter, 1971; Suzuki, 1971; Krinski, 1972) and has been used to obtain a number of important results on both the X-Y-Z model and the 8-vertex model (Baxter, 1971; Johnson, Krinsky and McCoy, 1973). In chapter 3 we have shown that the knowledge of *all* eigenvalues of the two-dimensional Ising model, which is a special case of the 8-vertex model (for a review see: Lieb and Wu, 1972), is required to find a closed-form expression for the free energy of the one-dimensional Ising model in a transverse field. For the more general 8-vertex model only the few largest (in absolute value) eigenvalues are known (Baxter, 1971, 1972a, 1972b; Johnson, Krinsky and McCoy, 1973). However it is possible to compute thermodynamic properties of the model (5.1) by solving a hierarchy of integral equations (Gaudin, 1971; Johnson and McCoy, 1972; Takahashi, 1973; Johnson, 1974; Johnson and Bonner 1980).

Considering the enormous amount of available results on the static properties of (5.1) and models related to it (Bonner et al., 1981) and the analogies mentioned above it is clear that one should not expect that use of the Trotter formula will offer an alternative computational technique for solving (5.1). Moreover the Bethe ansatz approach (Bethe, 1931; Yang and Yang, 1966), which has been the basis of most of the analytic results and numerical diagonalization of short chains, has been used to study spin *dynamics* (Des Cloizeau and Pearson, 1962; Müller, Beck and Bonner, 1979; Johnson and Bonner, 1980; Müller et al., 1981; Schneider and Stoll, 1981), an important aspect of the problem for which the Trotter-formula approach has not yet produced many useful results (some results can be found in Kalos, 1984).

One reason for studying path-sum representations of the one-dimensional model (5.1) is to gain insight in the properties of the Trotter approximations. Our main motivation is that in section 5.2 we shall use the local partitioning to calculate the thermodynamic properties of the two-dimensional spin-1/2 XY model. Having experience with the much simpler one-dimensional case is essential for a successful application of the path-summation approach to two-dimensional spin-1/2 models.

Now we briefly summarize the conclusions drawn from exact numerical calculations of short chains, more details can be found elsewhere (De Raedt, Lagendijk and Fivez, 1982). As the approximants to the partition function are the same for the checkerboard and real-space decomposition all thermodynamic functions are independent of the particular choice. However static spin-correlation functions certainly depend on the choice of break-up and although the real space partitioning breaks translational invariance (in the sense that for instance $\langle \sigma_1^z \sigma_2^z \rangle_m \neq \langle \sigma_2^z \sigma_3^z \rangle_m = \langle \sigma_3^z \sigma_4^z \rangle_m = \cdots \neq \langle \sigma_{M-1}^z \sigma_M^z \rangle_m$, see the discussion at the end of section 4.3), we find that real-space approximants to expectation values of translational invariant operators converge somewhat faster than those obtained from the checkerboard representation. The effect of breaking rotational invariance by decomposing the Hamiltonian into XY- and Ising models turns out to be very small except for a few specific correlation functions that do not seem to converge at all for the values of m considered. In table 5.1 we give typical results obtained from the checkerboard approximation to the thermodynamic functions of the one-dimensional XY model. We conclude that although the temperature is fairly low the convergence is quite good. Also note that the approximate free energy and energy both increase monotonically whereas the specific heat does not. For small m the approximation to the entropy $S_m = \beta(F_m - E_m)$ is negative.

The most important conclusion is that if we choose to work with the local split-up the rate of convergence of all relevant physical quantities is quite good. In general we find that if we increase the lattice size and keep all other parameters constant the relative error increases slowly. The ap-

Table 5.1
The approximations to the free energy per site (F_m^{CBD}), the energy per site (E_m^{CBD}) and the specific heat per site (C_m^{CBD}) of the spin-1/2 XY model ($J_3 = h = 0$) for a ring of $M = 6$ spins. The temperature $T/J_1 = 0.5$ corresponds to a very low temperature (the exact ground-state energy per site $E_g/M = -8/6 \approx 1.3333$, the exact thermal energy $E/M = -1.2978$). The numbers corresponding to $m = \infty$ have been obtained by numerical diagonalization of (5.1)

m	F_m^{CBD}	E_m^{CBD}	C_m^{CBD}
1	−1.7698	−1.9933	0.1067
2	−1.5744	−1.8570	1.0645
4	−1.4082	−1.4484	0.5587
8	−1.3656	−1.3367	0.2194
16	−1.3547	−1.3076	0.1456
32	−1.3520	−1.3002	0.1400
64	−1.3513	−1.2984	0.1419
128	−1.3511	−1.2980	0.1428
256	−1.3511	−1.2978	0.1432
1024	−1.3511	−1.2978	0.1433
∞	−1.3511	−1.2978	0.1434

proximations to the Heisenberg ferromagnet ($J_1 = J_3 = J > h = 0$) converge much faster than those to the XY or antiferromagnetic models. This is due to the fact that quantum fluctuations are much more important in the latter models than in the ferromagnet as can for instance be seen by comparing the ground-state wave functions. Another way of looking at this is to imagine what would happen with the checkerboard representation if we keep m finite and let $T \to 0$ (obviously this is not the correct procedure to obtain the ground-state properties). Then one easily sees that the ground state obtained from the approximation is the product of the ground states of the two-site blocks which in the case of the ferromagnet is also the true ground state of the infinite system. Thus even for $m = 1$ we will recover the exact ground-state energy (but not all other properties) of the Heisenberg ferromagnet and it is therefore understandable that for this case the convergence is better.

A notably complete discussion of the difficulties that are encountered when the Monte Carlo method is utilized to calculate the properties of the spin-1/2 chain has been given by Cullen and Landau (Cullen and Landau, 1983). They carried out simulations for the ferromagnetic and antiferromagnetic Heisenberg chain and the one-dimensional XY model by using the checkerboard representation (or the equivalent 6-vertex formulation). They combined simulation data of rings of 32 spins and $m = 1, \ldots, 8$ and finite-size scaling theory (Fisher and Ferdinand, 1967; Ferdinand and Fisher, 1969) to extrapolate the finite-m Monte Carlo data to $m = \infty$.

Implementing a Monte Carlo scheme for the checkerboard lattice is not as easy as simulating normal Ising models only because there are a number of configurations (8 out of 16 in the case of the general X-Y-Z model, 10 out of 16 in the case of the isotropic Heisenberg model) of each set of four interacting spins that have zero weight and are therefore forbidden. In terms of vertex model terminology we have to simulate a 6-vertex model and whether or not this model originates from mapping the spin-1/2 chain onto a vertex model, Monte Carlo simulations of 6-vertex models suffer from the inherent difficulty that it is hard to find good and efficient algorithms that generate, without bias, only allowed vertex configurations.

For the problem at hand this is most evident from (5.4) where we see that because we have to fulfill $\Pi_{i=1}^{M} S_{i,j}S_{i,j+1} = 1$ at all times it is not possible to use an algorithm that flips an arbitrary set of spins. For instance flipping of a single Ising spin, a procedure that is sufficient in the case of genuine Ising models, is not allowed. However we can readily construct two elementary spin-flip procedures from which all more complicated spin-flip algorithms can be built. Assuming that we start from an allowed configuration we can flip the set of spins $\{S_{i,1}, S_{i,2}, \ldots, S_{i,m}\}$ thereby changing the total magnetization $\Sigma_{i=1}^{M} S_{i,j}$ of the system. If we keep the total magnetization constant, the constraint $\Pi_{i=1}^{M} S_{i,j}S_{i,j+1} = 1$ is automatically satisfied. Thus the second type of elementary step is to flip two neighboring spins $S_{i,j}$ and $S_{i+1,j}$. It is not difficult to image that given an allowed configuration we can construct any other allowed configuration by a suitable combination of the two elementary spin-flip procedures.

For the 8-vertex model the resulting multi-spin flip algorithm would yield an efficient scheme to generate the vertex configurations and consequently this scheme can be used to simulate the X-Y-Z chain. However in the case of the isotropic Heisenberg model or XY model the checkerboard lattice (5.3) corresponds to a 6-vertex model and it turns out that it is difficult to devise an efficient algorithm that automatically takes account of the two extra zero-weight configurations. In their simulation work Cullen and Landau allow for the generation of the two forbidden states in the spin-flip procedure but do not accept these two states as new states. As the number of forbidden states is a fast increasing function of m, computer time used on waisted trials is a major bottle neck and in practice there seems to be little point of trying to go beyond $m = 8$ (Cullen and Landau, 1983).

Monte Carlo results for the energy of the Heisenberg ferromagnet (Cullen and Landau, 1983) are shown in fig. 5.3. As stated above the convergence of the approximate energy to the exact result is good. Finite-size scaling analysis of critical behavior ($T \to 0$) yields a correlation exponent $\nu = 1.02 \pm 5\%$ and a susceptibility exponent $\gamma = 1.32 \pm 10\%$ (Cullen and Landau, 1983). The value for γ disagrees with

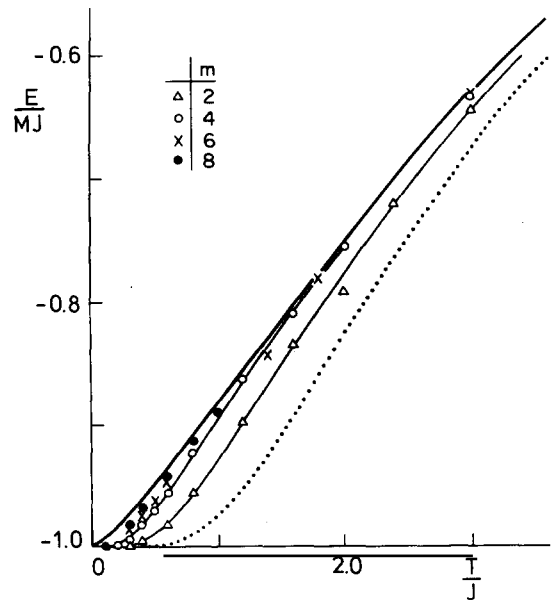

Fig. 5.3. Temperature dependence of the energy of the ferromagnetic Heisenberg ring of 32 spins. The bold solid line is obtained by extrapolating short chain diagonalization data (Bonner and Fisher, 1966), the dotted line is the exact $m = 1$ result (Suzuki, 1966). Other lines are guides to the eye. After Cullen and Landau, 1983.

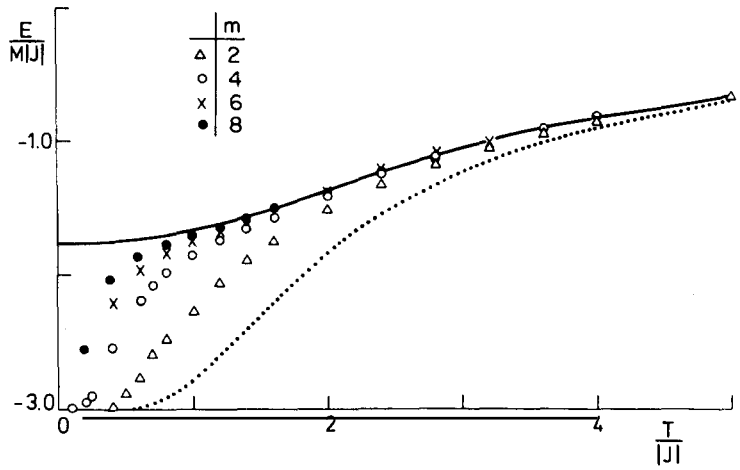

Fig. 5.4. Temperature dependence of the energy of the antiferromagnetic Heisenberg ring of 32 spins. The bold solid line is obtained by extrapolating short chain diagonalization data (Bonner and Fisher, 1966), the dotted line is the exact $m = 1$ result (Suzuki, 1966). Other lines are guides to the eye. After Cullen and Landau, 1983.

the estimate from high-temperature series analysis $\gamma = 1.66 \pm 10\%$ (Baker, Rushbrooke and Gilbert, 1964), the extrapolated short chain result $\gamma = 1.8$ (Bonner and Fisher, 1964) and a recent Monte Carlo estimate $\gamma = 1.75$ (Lyklema, 1983). A possible explanation for this disagreement is found by recalling that in general the critical exponents of the 6-vertex model depend on the interactions (Baxter, 1971, 1972a) which in our case is tantamount to the statement that the m-dependence of the exponents may be complicated.

Simulation results for the energy of an antiferromagnetic Heisenberg ring of 32 spins are depicted in fig. 5.4. Unlike for the ferromagnet at $T = 0$, quantum effects are very strong as can be seen immediately by comparing the classical Néel ground-state energy per site $E_{\text{Néel}}/JM = 1$ with the exact ground-state energy per site $E_0/JM = 4 \ln 2 - 1 \approx 1.76$ (Hulthén, 1938). Thus one might anticipate to find much slower convergence as a function of m but simulation data for the energy (Cullen and Landau, 1983) demonstrate that for $m = 8$ it is possible to obtain good estimates for the energy as long as $T/|J| > 0.5$. For a ring of 8 spins at $T/|J| = 0.5$ relative difference between ground-state energy and thermal energy is approximately 0.6% and it is therefore not unreasonable to say that $T/|J| = 0.5$ corresponds to a low temperature. A detailed discussion of simulation data of the specific heat and static susceptibility of the Heisenberg model and of similar results for the XY chain can be found in the original paper (Cullen and Landau, 1983).

5.2. Two-dimensional spin-1/2 XY model

Here we employ the path-summation approach to attack a problem which so far has withstood all attempts to calculate the temperature-dependent properties by "convential" techniques such as quantum versions of the renormalization group methods (Betts and Plischke, 1976; Rogiers and Dekeyser, 1976; Dekeyser et al., 1978; Stella and Toigo, 1978; Takano and Suzuki, 1981; Tatsumi, 1981) and series-expansion analysis (Betts, 1974). There are two different major issues, namely critical behavior and low temperature properties. It is known for some time that the two-dimensional spin-1/2 XY model ($J_3 = h = 0$ in (5.1)) belongs to the class of models for which it has been shown rigorously that they will

not exhibit spontaneous magnetization (Mermin and Wagner, 1966). On the other hand for classical versions of the model (planar rotator models) it is now well established that although there is no spontaneous magnetization there is a phase transition at non-zero temperature (Stanley and Kaplan, 1966; Wegner, 1967; Betts, 1974). This phase transition is attributed to the existence of topological excitations in the system (Berezinskii, 1970; Kosterlitz and Thouless, 1973) and has been studied extensively (Villain, 1975; Kosterlitz, 1974; José et al., 1977; Zittartz, 1978a, 1978b, 1978c; Tobochnik and Chester, 1979; Fröhlich and Spencer, 1981). The physical picture is that apart from the usual spin waves there are also bound topological defects (spin vortices) at sufficiently low temperature and it is the unbinding of pairs of defects with increasing temperature that causes a phase transition to take place. Whether or not the same physical picture can be extended to the 2D spin-1/2 XY model is an open question. It is well-known that in two dimensions non-universal behavior is not unusual (Kogut, 1979). In particular it has been suggested that for this type of model the universality class depends on the spin (Betts, 1974). Thus results for the *classical* 2D *planar* model should not be taken as a criterion for the calculations of the 2D spin-1/2 XY model. At a very low temperature the properties of the quantum model do not even resemble those of the classical equivalent since the ground state of the quantum model is not a fully ordered state but rather a linear combination of states with magnetization zero (Mattis, 1979). The ground-state properties have been studied by small-lattice diagonalization (Oitmaa and Betts, 1978; Betts, Salevsky and Rogiers, 1981; Kelland, Betts and Oitmaa, 1981; Betts and Kelland, 1983) and by variational methods (Suzuki and Miyashita, 1978). In the low-temperature regime the path-summation approach can be used, but only with much greater computational effort.

For the path-summation approach to the two-dimensional spin-1/2 XY model we follow two different strategies. Recalling the common lore that for this model quantum effects should not be very important in the temperature region where critical behavior is expected to occur we may hope that even $m = 1$ approximations will give a reasonable description. Thus we will first confine ourselves to a special class of $m = 1$ approximations and show that they can be solved rigorously (Lagendijk and De Raedt, 1982). Then we make use of the tricks learned by playing around with the one-dimensional models to construct $m \geq 1$ representations that allow an efficient implementation of the Monte Carlo algorithm.

In order to prove that it is possible to solve particular $m = 1$ approximations we make a graphical construction that maps these representations into a rigorously solvable two-dimensional lattice model, a staggered 8-vertex model (Hsue, Lin and Wu, 1975). The construction parallels the one used to map the checkerboard representation of the one-dimensional model onto an 8-vertex model (Barma and Shastry, 1978). For the local two-site partitioning the $m = 1$ approximation reads

$$Z_1 \equiv \mathrm{Tr} \prod_{\langle ij \rangle} \exp(-\beta H_{i,j}) . \tag{5.6}$$

As each spin of the square lattice interacts with four neighbors it is clear that by inserting resolutions of the identity we will end up with $4L^2$ Ising spins (L is the linear size of the square). Thus as shown in fig. 5.5 we may place the four Ising spin variables on the corners of a small square that contains the corresponding quantum spin. Now we have to find a way to order the two-site operators such that all bonds are taken into account. Let us start by considering the two-site operator acting on the spins on sites 1 and 2 (see fig. 5.5). If we label the Ising spin variable by the number of the site enclosed by the elementary square and one of the letters a, b, c, d it is clear that the matrix element for the bond 1–2 can be denoted by $\langle S_{1,a} S_{2,a'} | \exp(-\beta H_{1,2}) | S_{1,b} S_{2,b'} \rangle$. Now because we are inserting resolutions of the identity it is obvious that the next time we encounter an interaction in which for instance the spin at site

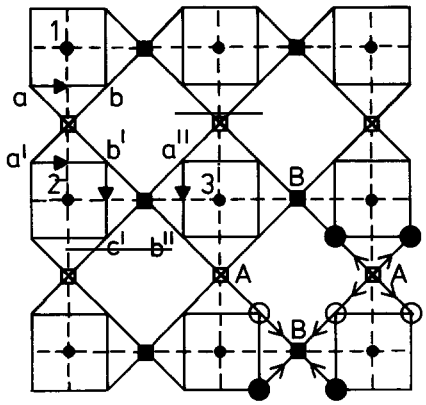

Fig. 5.5. The effective Ising lattice model for the $m = 1$ approximation to the partition function of the two-dimensional spin-1/2 XY model. The original lattice is given by the dashed lines. The Ising spins live on the corners of the squares that surround the original lattice points. The equivalent staggered 8-vertex lattice consists of sublattices A and B.

2 participates, we have to use either $S_{2,a'}$ or $S_{2,b'}$ in the bra respectively ket of the corresponding matrix element. For example, taking the interaction 2–3 into account yields the matrix element $\langle S_{2,b'} S_{3,a''} | \exp(-\beta H_{2,3}) | S_{2,c'} S_{3,b''} \rangle$.

The notation used so far becomes very clumsy after taking a few bonds into account but one does not need to keep track of all the indices. To this end we introduce the convention that if we add a matrix element to the product under construction, we draw two arrows, pointing to the Ising spins that appear in the ket of the matrix element, on the edges of the squares that cross the bond (dashed lines in fig. 5.5). Now the rule of successively using the Ising spins of a particular square is replaced by the requirement that the handedness in which the arrows run over the squares cannot be changed. Note that once one of the Ising spins of a square is used in a bra or ket the handedness on the square is fixed. Using this notation we can walk over the lattice and add more and more exponents of two-site Hamiltonians to the list such that at the end there are exactly four arrows on each square (in the case of periodic boundary conditions). Then all Ising spins have been used twice (once in a bra and once in a ket) and the path-sum representation has been found. It is not difficult to convince oneself that this construction can be made for many (but not all) different orderings under the condition that L is even.

As for the one-dimensional case we turn the Ising model into a vertex model by drawing diagonals in the squares between (instead of on) the elementary squares on which the Ising spins live, using the same convention for putting arrows on the lines through a corner and attaching the appropriate weights to the vertex configurations. From figs. 5.2 and 5.5 it follows immediately that the resulting vertex model is a staggered 8-vertex (SEV) model with weights given by

$$\omega_1 = \omega_1' = \omega_2 = \omega_2' = 1\,, \tag{5.7a}$$

$$\omega_3 = \omega_3' = \omega_4 = \omega_4' = \sinh 2\beta J_1\,, \tag{5.7b}$$

$$\omega_5 = \omega_7' = \omega_6 = \omega_8' = \cosh 2\beta J_1\,, \tag{5.7c}$$

$$\omega_7 = \omega_5' = \omega_8 = \omega_6' = 0\,, \tag{5.7d}$$

where ω_i (ω_i') are the weights on the sublattice A (B).

The general solution for the staggered 8-vertex model is not known but a rigorous closed form expression for the free energy can be derived under the assumption that the vertex weights obey the free-fermion condition (Hsue, Lin and Wu, 1975). For the 2D spin-1/2 XY model this happens to be the case and the zero field result for the approximation to the free energy per site reads

$$F_1^{\mathrm{SEV}}/M = -\frac{1}{8\pi^2\beta}\int_0^{2\pi}\int_0^{2\pi} \mathrm{d}\phi\,\mathrm{d}\theta\ \ln[4(1+y^2)^2 - 4y^2(\cos\theta+\cos\phi)^2]\,, \tag{5.8}$$

where $y = \sinh 2\beta J_1$. From (5.8) we already see that the critical point must be at $y = 1$ ($T_1^c/J_1 \approx 2.27$) since the integrand diverges for $\theta = \phi = 0$. A more detailed analysis reveals that the specific heat has a logarithmic divergence at the Ising model critical temperature $\sinh(2J_1/T_1^c) = 1$. As the free-fermion condition remains satisfied if we add a longitudinal field $\boldsymbol{h} = h\boldsymbol{e}_z$ we can examine the behavior of the magnetization and the longitudinal static susceptibility and we find that there is no spontaneous magnetization nor a divergence in the longitudinal susceptibility (Lagendijk and De Raedt, 1982). Thus the rigorous solution of the simplest approximation to the partition function of the two-dimensional spin-1/2 XY model yields results which are not in conflict with all known rigorous results on the model.

Of course on the basis of (5.8) alone it is not possible to answer the question whether the obtained critical behavior is not an artifact of choosing a particular ordering of the exponential operators or due to taking $m = 1$. We now examine the problem of extending the calculation to $m \geq 1$ and other orderings. Extending the staggered 8-vertex formulation to $m > 1$ is possible but it is certainly not the most promising approach if we have to use the Monte Carlo simulation method. As two out of eight vertex configurations are forbidden we inevitably face the same difficulties as those encountered in simulations of the one-dimensional spin-1/2 models (Cullen and Landau, 1983). In addition we would have to find an efficient scheme to deal with the complicated three-dimensional lattice structure. Thus we consider extensions of staggered 8-vertex formulations as being impractical.

Evidently, the major stumbling block is the large number of zero's in the two-site interaction matrix (5.3). In the case of the XY model (in zero external field $\boldsymbol{h} = 0$) there is a representation of the spin states such that eight of the sixteen elements are strictly positive. Applying two successive cyclic permutations of the spin components we find that the XY model turns into an XZ model

$$H = -J\sum_{\langle ij\rangle}(\sigma_i^x\sigma_j^x + \sigma_i^z\sigma_j^z)\,. \tag{5.9}$$

Since a cyclic permutation is a unitary transformation the partition function does not change. It is straightforward to show that for the XZ model the two-site interaction matrix is given by

$$\begin{aligned} T(S_i, S_j; \bar{S}_i, \bar{S}_j) &= \langle S_i S_j|\exp\left[\frac{\beta J}{m}(\sigma_i^x\sigma_j^x + \sigma_i^z\sigma_j^z)\right]|\bar{S}_i\bar{S}_j\rangle \\ &= \delta_{S_iS_j,\,\bar{S}_i\bar{S}_j}(\tfrac{1}{2}\sinh 2K_m)^{1/2}\exp\left(K_m S_i\bar{S}_j + \frac{\beta J}{m}S_iS_j\right), \end{aligned} \tag{5.10}$$

where $K_m = \frac{1}{2}\ln\coth(\beta J/m)$. From (5.10) it is already evident that we have found a representation that will not suffer from an fundamental limitation. Indeed as pointed out above, simulating an 8-vertex model is more complicated than simulating an Ising model but not impossible and furthermore all eight

non-zero matrix elements are strictly positive. We might wonder whether for the more general Hamiltonian there might exist representations that have these special properties. The answer is certainly no because for the isotropic Heisenberg model we can rotate the spins as much as we want, we always end up with a coupling matrix that has ten zeros in it.

The remaining problem is to split up the Hamiltonian such that the labor involved in keeping track of all two-site interactions is not too exhaustive. To reduce the computation time spent in calculating indices of interacting spins we have used a particularly simple partitioning. We decompose the Hamiltonian as $H = H_h + H_v$ where $H_{h(v)}$ is a sum of non-interacting horizontal (vertical) XZ chains. Then we can either choose the checkerboard or real-space partitioning to break-up the XZ chains into two-site blocks. In this last step we do the same as for the one-dimensional models, i.e. we sum out intermediate states. Thus instead of $4mL^2$ Ising-spin variables (4 per quantum spin because each quantum spin interacts with four other quantum spins) we can reduce the total number of variables by a factor of two.

Following the prescriptions outlined above it is nothing but a tedious exercise in manipulating indices to rewrite the mth approximant to the partition function of the two-dimensional spin-1/2 XZ model as a three-dimensional Ising model with complicated many-spin interactions. We find (De Raedt et al., 1984)

$$Z_m = c(\sinh 2K_m)^{mL^2} \sum_{\{S_{i,j}^{(k)}\}}{}' \sum_{\{\bar{S}_{i,j}^{(k)}\}}{}' \prod_{k=1}^{m} \prod_{j=1}^{L} \prod_{i=1}^{L} h(j,k)\, v(i,k)\,, \tag{5.11}$$

where c is an unimportant numerical constant. The prime on the summation symbols mean that the sums over the Ising spin variables are restricted by $2mL-1$ constraints that can be written as

$$\bar{S}_{1,1}^{(k)} = S_{1,1}^{(k)} \prod_{i=2}^{L} S_{i,1}^{(k)} S_{i,1}^{(k+1)} \prod_{i=2}^{L} \prod_{j=2}^{L} \bar{S}_{i,j}^{(k)} S_{i,j}^{(k+1)}\,, \tag{5.12a}$$

$$\bar{S}_{1,j}^{(k)} = S_{1,j}^{(k)} \prod_{i=1}^{L} S_{i,j}^{(k)} \bar{S}_{i,j}^{(k)}; \qquad j>1\,, \tag{5.12b}$$

$$\bar{S}_{i,1}^{(k)} = S_{i,1}^{(k)} \prod_{j=1}^{L} \bar{S}_{i,j}^{(k)} S_{i,j}^{(k+1)}; \qquad i>1\,, \tag{5.12c}$$

$$S_{1,1}^{(k)} = S_{1,1}^{(1)} \prod_{i=2}^{L} \prod_{j=1}^{L} S_{i,j}^{(1)} S_{i,j}^{(k)} \prod_{j=2}^{L} S_{1,j}^{(1)} S_{1,j}^{(k)}; \qquad k>1\,. \tag{5.12d}$$

For periodic boundary conditions and real-space split-up of the XZ chains $h(j\,k)$ is given by

$$\begin{aligned} h(j,k) = \cosh\Big[K_m\big(&S_{1,j}^{(k)}\bar{S}_{1,j}^{(k)} + \cdots + S_{1,j}^{(k)}\cdots S_{L,j}^{(k)}\bar{S}_{1,j}^{(k)}\cdots\bar{S}_{L,j}^{(k)}\big) \\ &+ \frac{\beta J}{m}\big(S_{1,j}^{(k)}S_{2,j}^{(k)}S_{3,j}^{(k)}\bar{S}_{1,j}^{(k)} + \cdots + S_{1,j}^{(k)}\cdots S_{L,j}^{(k)}\bar{S}_{1,j}^{(k)}\cdots\bar{S}_{L-2,j}^{(k)}\big)\Big] \\ &\times \exp\Big[\frac{\beta J}{m}\big(S_{1,j}^{(k)}S_{2,j}^{(k)} + S_{1,j}^{(k)}\cdots S_{L,j}^{(k)}\bar{S}_{2,j}^{(k)}\cdots S_{L-1,j}^{(k)}\big)\Big]\,, \end{aligned} \tag{5.13}$$

whereas for periodic boundary conditions and checkerboard break-up

$$h(j,k) = \cosh[K_m(S_{1,j}^{(k)}\bar{S}_{1,j}^{(k)} + \cdots + S_{1,j}^{(k)}\cdots S_{L,j}^{(k)}\bar{S}_{1,j}^{(k)}\cdots\bar{S}_{L,j}^{(k)})]$$
$$\times \exp\left[\frac{\beta J}{m}(S_{1,j}^{(k)}S_{2,j}^{(k)} + S_{3,j}^{(k)}S_{4,j}^{(k)} + \cdots + S_{L-1,j}^{(k)}S_{L,j}^{(k)}\right]$$
$$\times \exp\left[\frac{\beta J}{m}(\bar{S}_{2,j}^{(k)}\bar{S}_{3,j}^{(k)} + \bar{S}_{4,j}^{(k)}\bar{S}_{5,j}^{(k)} + \cdots + \bar{S}_{L,j}^{(k)}\bar{S}_{1,j}^{(k)}\right]. \tag{5.14}$$

The corresponding expression for $v(i,k)$ is obtained from $h(j,k)$ by replacing the symbols $S_{i,j}^{(k)} \to \bar{S}_{i,l}^{(k)}$ and $\bar{S}_{i,j}^{(k)} \to S_{i,l}^{(k+1)}$. The 3D model (5.11) has complicated many-spin interactions and a coupling that depends on the lattice size in the Trotter direction (labeled by the superscript k).

In the Monte Carlo procedure for model (5.11) we can flip any single Ising spin that does not appear in the left-hand side of (5.12) and (5.12) will tell us which other spins we have to change. Although several spins flip at the same time we still call this a single spin-flip procedure. Since we change several spins at each step one might think that because of the highly non-local interactions in (5.13) or (5.14), one Monte Carlo step will take much computing time. This would certainly be the case if we would use one computer storage unit (i.e. a word) per spin variable because we had to make several loops to update spins and calculate the transition probability. As we are simulating an Ising model we may as well use one bit to store an Ising spin and pack each row and column into one word. Then updating many spins in one row or column means that we can use masking operations (AND, OR and EXLUSIVE OR) on words. Boolean operations on integer variables are not supported by standard Fortran 77 compilers but most Fortran compilers (on DEC, CDC and IBM machines) allow for this kind of operations and produce an extremely efficient code. The use of this technique is necessary to get acceptable program performance but it is clear that simulation of the 2D XY model will take much more time than simulating a normal Ising model (with nearest-neighbor interactions) of the same size since each Ising spin "interacts" with other Ising spins through a complicated, non-local action.

We have tested the single spin-flip algorithm by simulating small systems (up to 3×3) and have compared the data with results obtained from exact enumerations and diagonalizations and found excellent agreement. In going to larger lattices (6×6) we observed that for high temperatures, the data deviated systematically from the high-temperature series results (Betts and Plischke, 1976). The Monte Carlo algorithm itself also signalled that something was going wrong because the acceptance rate became extremely small. From (5.14) or (5.15) it follows that if K_m is large, and this happens if $T\to\infty$ or $m\to\infty$, changing spins without keeping the prefactor K_m constant is very difficult. A simple way out of this problem is to flip the set of spins $\{S_{i,j}^{(1)},\ldots,S_{i,j}^{(m)},\bar{S}_{i,j}^{(1)},\ldots,\bar{S}_{i,j}^{(m)}\}$. Such a multi-spin flip step keeps (5.12) intact and also requires less computing time.

Combination of the two spin-flip procedures yields a Monte Carlo scheme that reproduces all known (i.e. small systems, high-T) results of the 2D spin-1/2 XY model. In practice we found that the "classical" multi-spin flip step was the most important in the sense that it was sufficient to keep the ratio, single spin-flips over multi spin-flips, low (10%). All data presented in this paper have been obtained from at least two independent runs of 10 000 Monte Carlo steps per Ising spin each. For the largest lattices that we have studied it was necessary to use a CYBER 205 (see De Raedt, De Raedt and Lagendijk, 1984). Simulation of an $L=32$, $m=8$ ($32\times 32\times 8$, 16384 Ising spins) system takes 3 hours of CPU time on a CYBER 205.

During a simulation of the 2D spin-1/2 XY model, we want to sample the energy, specific heat and spin correlation functions that change drastically if the system undergoes a phase transition (assuming that there is one). The static susceptibilities are such correlation functions but as their calculation requires a double summation over a big 3D lattice (see (2.10c)) and this is a time consuming computation, we decided to look for an alternative way to characterize the phase transition. In particular we have found it expedient to study a correlation function that gives us information about the degree of disorder in the system. It has been noted previously that for models in which there occurs a phase transition without long range order (such as the 2D XY model) it is more convenient to study *disorder* parameters than to look for quantities that describe the degree of order (Kogut, 1979). Moreover in models which are self-dual (such as the 2D Ising model) disorder and order are intimately related to each other.

One of the correlation functions that we have measured in our simulations is

$$D = L^{-2} \sum_{\text{plaquettes}} \langle (1 - \sigma_i^x \sigma_k^x - \sigma_i^z \sigma_k^z)(1 - \sigma_j^x \sigma_l^x - \sigma_j^z \sigma_l^z) \rangle , \tag{5.15}$$

where $i(j)$ and $k(l)$ label the opposite corners of an elementary square. In studies of the classical analogon, the planar rotator model, correlation function (5.15) gives information about the occurrence of vortex-like excitations (Swendsen, 1982). We will call D a vortex detector although no simple intuitive picture is available for vortex-like excitations in the quantum XY system. Results for two-spin correlation functions and the corresponding structure factors can be found in (De Raedt, De Raedt and Lagendijk, 1984).

Guided by the rigorous $m = 1$ result (for a different ordering) that there is a phase transition at $T_1^c/J \approx 2.27$, we first perform simulations of systems of different size for $m = 1$. In fig. 5.6 we show $m = 1$ results for the energy per site, obtained from the SEV, checkerboard and real-space representation. We see that the results for the three different orderings are almost the same unless the temperature $T/J < 2.25$. Also the size dependence (for $L \geq 6$) is rather weak. At low temperature $T/J = 1$ the approximate energy E_1 is much lower than the accurate estimate $E_g/JM \approx -2.2$ (Pearson, 1977) but since the XY real-space approach yields energies which are systematically higher than those of the checkerboard approximation we decided to use the real-space approximation in most of the

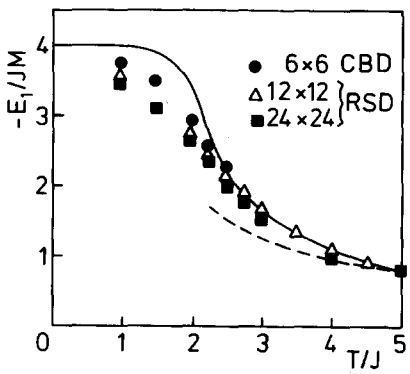

Fig. 5.6. Energy per site of the 2D spin-1/2 XY model in the $m = 1$ approximation. Solid line: rigorous staggered 8-vertex solution, full circles: checkerboard decomposition (CBD), full squares and open triangles: real-space decomposition (RSD) and broken line: high-temperature expansion.

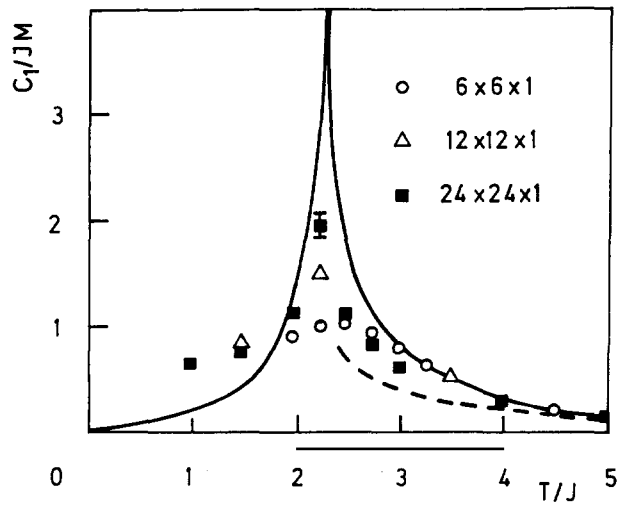

Fig. 5.7. Specific heat per site of the 2D spin-1/2 XY model in the $m = 1$ approximation. Solid line: rigorous staggered 8-vertex solution, open circles, full squares and open triangles: real-space decomposition: (RSD).

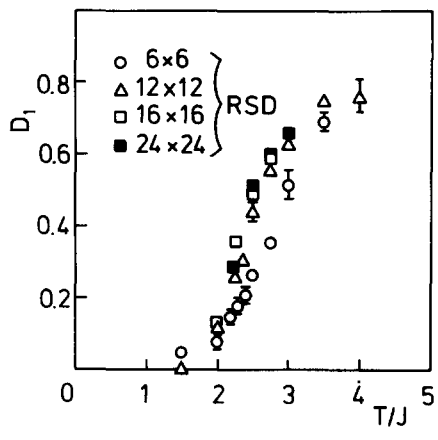

Fig. 5.8. The vortex correlation function in the $m = 1$ real-space approximation.

simulations. In fig. 5.7 we depict simulation data for the $m = 1$ approximation in most of the simulations. In fig. 5.7 we depict simulation data for the $m = 1$ approximant (C_1) to the specific heat. At $K \approx K_1^c$ the specific heat exhibits a maximum that grows slowly with the lattice size. This is not inconsistent with the exact $m = 1$ solution which predicts a logarithmically divergent specific heat (Lagendijk and De Raedt, 1982). From the $m = 1$ simulation data we conclude that the phase transition in the $m = 1$ SEV approximation is not the result of choosing a particular ordering. The first attempt to simulate the 2D XY model was reported by Suzuki et al. (Suzuki et al., 1977) who only studied the $m = 1$ case. Their simulation data for the specific heat disagree with ours and is also in qualitative disagreement with the rigorous $m = 1$ solution. The temperature dependence of the $m = 1$ vortex-detector is shown in fig. 5.8. It is clear that D_1 changes rapidly if T approaches T_1^c indicating that the system might exhibit a peculiar kind of disorder if $T \geq T_1^c$. We find that for $L \geq 8$ the size dependence of D_1 is small.

The power of the approach is that by increasing m we can improve the approximation systematically. From the $m > 1$ data shown in fig. 5.9 we conclude that for $T/J > 2$ the specific heat, which from the point of view of convergence is the most difficult quantity to calculate, depends weakly on the particular value of m. Except for $2 \leq T/J \leq 2.5$ the size dependence of the energy per site and specific heat per site is small. More convincing evidence that the convergence of the energy and specific heat is very good is given in fig. 5.10. In the critical region ($T/J \approx 2.27$) the m-dependence of physical quantities is very weak. Therefore it might be tempting to assume that to a good approximation the critical properties of the 2D spin-1/2 XY model are that of the $m = 1$ representation. Although our numerical data are not inconsistent with this assumption we take the point of view that, as in the 1D case, the subtitle β/m dependence of the approximants could change the critical behavior. Furthermore it is obvious that it is impossible to prove or disprove by means of Monte Carlo data that some physical quantity is continuous or divergent. Simulation data for the approximant D_m are shown in fig. 5.11 As for the energy and specific heat, in the critical region the m-dependence of D_m is very weak.

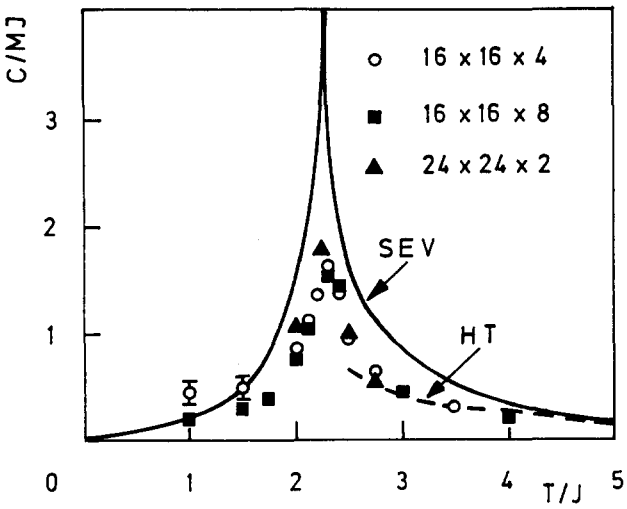

Fig. 5.9. Specific heat per site of the 2D spin-1/2 XY. Solid line: rigorous staggered 8-vertex solution; broken line: high-temperature expansion. Comparison with the data of fig. 5.7 shows that in the critical region the m-dependence is weak.

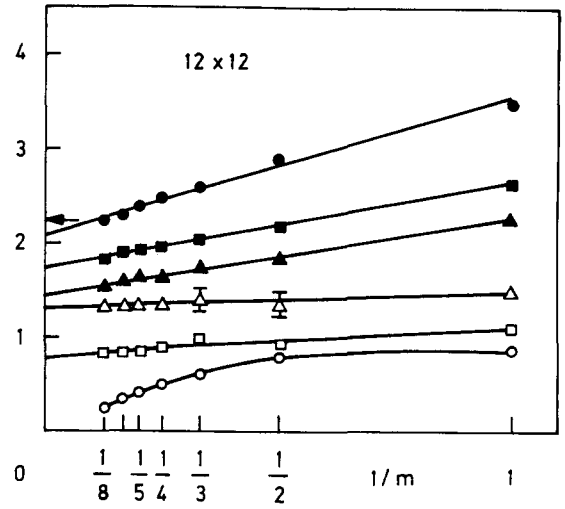

Fig. 5.10. Minus the energy per site (full symbols) and specific heat per site of the 2D spin-1/2 XY model (open symbols) as a function of $1/m$. Circles: $T/J = 1$, squares: $T/J = 2$ and triangles $T/J = 2.25$. The arrow gives the Pearsons estimate of the ground-state energy.

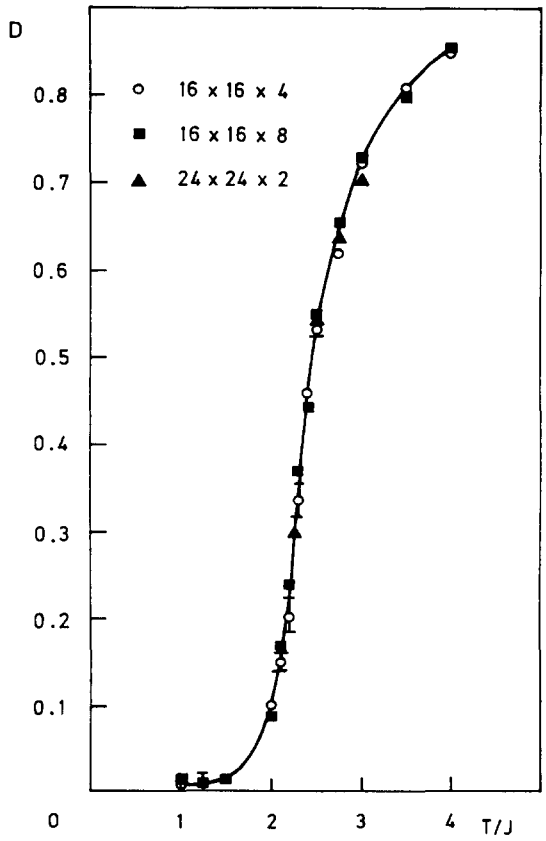

Fig. 5.11. The vortex correlation function, which measures the degree of disorder in the 2D spin-1/2 XY model, for lattices of different size and different m. In the critical region the m-dependence of this correlation function is very weak.

To demonstrate the versatility of the method we have also carried out simulations for systems with $N = 8$, 16 and $1 \leq m \leq 32$ for temperatures down to $T/J = 0.5$. Within the statistical accuracy of the simulation data (which for the energy is less than one percent), we find that for $0.5 \leq T/J \leq 1.5$, the energy is constant. The extrapolated ground-state energy per site ≈ -2.2 which is in good agreement with estimates given by Pearson (Pearson, 1977).

5.3. *Handscomb's method: a short introduction*

This method is often compared with the path-summation approach and for this reason we will discuss it shortly. A Monte Carlo method that exploits the special properties of the spin-1/2 algebra to considerable extent has been proposed by Handscomb (Handscomb, 1962, 1964; Hammersley and Handscomb, 1964). In general we start from the Hamiltonian

$$H = H_0 + \sum_{i=1}^{M} H_i , \tag{5.16}$$

and assume that $[H_0, \Sigma_{i=1}^{M} H_i] = 0$. Then we expand $\exp(-\beta \Sigma_{i=1}^{M} H_i)$ in a Taylor series and obtain for the partition function

$$Z = \sum_{n=0}^{\infty} \sum_{\{S_n\}} \frac{(-\beta)^n}{n!} \operatorname{Tr} e^{-\beta H_0} H_{i_1} H_{i_2} \cdots H_{i_n} , \tag{5.17}$$

where S_n denotes a sequence of the n indices $i_1, \ldots, i_n$; $1 \leq i_j \leq M$. The thermal expectation value of an operator A can then be written as

$$\langle A \rangle = \sum_{n=0}^{\infty} \sum_{\{S_n\}} A(S_n)\, \pi(S_n) , \tag{5.18a}$$

where

$$A(S_n) = \frac{\operatorname{Tr} e^{-\beta H_0} A H_{i_1} H_{i_2} \cdots H_{i_n}}{\operatorname{Tr} e^{-\beta H_0} H_{i_1} H_{i_2} \cdots H_{i_n}} \tag{5.18b}$$

and

$$\pi(S_n) = \frac{1}{Z} \frac{(-\beta)^n}{n!} \operatorname{Tr} e^{-\beta H_0} H_{i_1} H_{i_2} \cdots H_{i_n} . \tag{5.18c}$$

Supposing that for each sequence S_n the traces in (5.18) are easy to evaluate we may define the sample space as the set of all sequences S_n and interpret $\pi(S_n)$ as the probability for selection of a particular S_n. Evidently if $\pi(S_n)$ is to be a probability we must have $Z < \infty$ and $\forall S_n : \pi(S_n) > 0$. Then we can utilize the Monte Carlo technique to sample the abstract space of sequences S_n and measure physical properties by recording the quantities $A(S_n)$. The Markov chain of states S_n is generated by adding or removing an index k to the current list S_n. Specializing to the spin-1/2 chain (extension to two- or

three-dimensional models is straightforward) we take $H_0 = -h \Sigma_i \sigma_i^z$ and $H_i = \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1}$ and it becomes evident that on account of the low connectivity of the possible transitions it is not easy to design, add and remove schemes that give sufficient coverage of the sample space (Handscomb, 1962; Lyklema, 1982). If we enlarge the current set $S_n \to S_n k$ we have to choose k such that $\mathrm{Tr}\, e^{-\beta H_0} H_{i_1} \cdots H_{i_n} H_k \neq 0$ and a similar restriction appears if we shorten the list S_n. Lyklema (Lyklema, 1982) has extended the method by allowing the change in the current list S_n to occur everywhere in the list and not only at the end of it.

A remarkable feature of Handscomb's approach is that the dimensionality of the lattice does not appear explicitly in the Monte Carlo scheme and one might at first sight think that simulating a three-dimensional model takes no more time than simulating a one-dimensional system but this is of course not the case. To examine this point further it is sufficient to consider the expression for the zero field energy (Handscomb, 1964)

$$E/M = Jd - 2T\bar{n}, \tag{5.19}$$

where $\bar{n}$ is the average value of the length of the sequence S_n. From (5.19) we conclude that if $T \to 0$ we must have $\bar{n} = Jd/T + \mathcal{O}(1)$, $T \to 0$ in order to recover the correct ground-state energy of the Heisenberg ferromagnet $E_0/M = -Jd$. As the number of operations required to update the list S_n grows proportional with n (Handscomb, 1964) we conclude that the computation time increases with the dimensionality d and that when $T \to 0$ this method (as well as the Trotter formula approach) will certainly become impractical.

Handscomb's method has been tested on the Heisenberg ferromagnet (Lyklema, 1982, 1983; Chakravarty and Stein, 1982) and yields good results for the energy, specific heat and static susceptibility. Direct application of this approach to the antiferromagnetic Heisenberg model is not possible. Using the trick that solved the sign problem for one-dimensional fermions (De Raedt and Lagendijk, 1981a) an attempt has been made (Lyklema, 1982) to perform simulations for the one-dimensional antiferromagnet but this approach only yields reasonable answers at very high temperature $T/|J| > 3$ (Chakravarty and Stein, 1982) probably because there is no sound theoretical basis to assume that the trick will work. A non-trivial extension of Handscomb's method is required to perform simulations for the exactly solvable 1D XY model (Chakravarty and Stein, 1982). For relatively high temperatures $T/|J| \geq 1.4$ simulation data for the energy and specific heat of the XY chain are in good agreement with exact results.

In summary, Handscomb's method is a powerful method to study quantum ferromagnets but as yet it is not clear that the method is flexible enough to allow extensions to a broader class of unsolvable quantum lattice problems. To get more insight in the power of the method it would be helpful if more technical details would be published.

6. Systems coupled to bosons

6.1. General path-integral method

In this section we consider a system (H_1) which is coupled to a bath of harmonic oscillators (H_2) whereby the interaction of the system with the bosons is described by an operator (H_3) that is linear in the boson field. The general Hamiltonian for the coupled system takes the form

$$H = H_1 + H_2 + H_3\,, \tag{6.1a}$$

$$H_2 = \sum_{i=1}^{M} \frac{p_i^2}{2m_0}\,, \tag{6.1b}$$

$$H_3 = \sum_{i,\,i'=1}^{M} \left(\frac{m_0}{2} x_i A_{i,i'} x_{i'} + x_i B_{i,i'} Y_{i'}\right), \tag{6.1c}$$

where x_i is the coordinate and p_i the momentum of the oscillator at site i, m_0 is the mass of the oscillators, $A_{i,i'} = A_{|i-i'|}$ is the harmonic potential and $B_{i,i'} = B_{|i-i'|}$ is a coupling matrix. Since most of the analytic manipulations can be carried out without knowing the specific form of H_1 or the meaning of the operator Y_i we defer the discussion of the implications of choosing a particular H_1 and Y_i to the following sections where we look at some examples in more detail.

The main purpose of the rest of this section is to show how to make optimal use of the fact that Hamiltonian (6.1) is quadratic in both p_i and x_i. As demonstrated by Feynman in his path-integral treatment of the Fröhlich polaron (Feynman, 1955), a model which is of form (6.1), it is possible to perform the trace operation over all oscillator degrees of freedom analytically. Here we will closely follow the ideas of Feynman but since we are mainly interested in lattice models we will start from the Trotter formula in order to obtain a properly discretized version of the path integral. We decompose the Hamiltonian as already suggested by (6.1a) and work in the coordinate representation of the boson field and, for the system H_1, in the basis that diagonalizes the operators $Y_i (Y_i|y_i\rangle = y_i|y_i\rangle)$. The corresponding Trotter approximation reads

$$Z_m = \mathrm{Tr}[\exp(-\beta H_1/m)\exp(-\beta H_2/m)\exp(-\beta H_3/m)]^m \tag{6.2a}$$

$$= c\left(\frac{\beta}{m}\right)^{-mM/2} \sum_{\{y_{i,j}\}} \int \prod_{i=1}^{M}\prod_{j=1}^{m} \mathrm{d}x_{i,j} \exp[-S(\{x_{i,j}\},\{y_{i,j}\})]$$

$$\times \langle y_{1,j}\cdots y_{M,j}|\exp(-\beta H_1/m)|y_{1,j+1}\cdots y_{M,j+1}\rangle\,, \tag{6.2b}$$

where

$$S(\{x_{i,j}\},\{y_{i,j}\}) = \frac{mm_0}{2\beta}\sum_{i=1}^{M}\sum_{j=1}^{m}(x_{i,j}-x_{i,j+1})^2 + \frac{m_0\beta}{2m}\sum_{i,\,i'=1}^{M}\sum_{j=1}^{m}\left(x_{i,j}A_{i,i'}x_{i',j} + \frac{2}{m_0}x_{i,j}B_{i,i'}y_{i',j}\right) \tag{6.2c}$$

and c is an unimportant constant. Obviously (6.2c) is a quadratic function of the $x_{i,j}$ and a convenient way to evaluate the Gaussians integrals in (6.2) is to use

$$x_{n,j} = (mM)^{-1/2}\sum_{q=1}^{m}\sum_{p=1}^{M} v_{p,q}\exp\left(\frac{2\pi ijq}{m} + \frac{2\pi inp}{M}\right), \tag{6.3a}$$

$$y_{n,j} = (mM)^{-1/2}\sum_{q=1}^{m}\sum_{p=1}^{M} w_{p,q}\exp\left(\frac{2\pi ijq}{m} + \frac{2\pi inp}{M}\right), \tag{6.3b}$$

and to substitute this Fourier representation in (6.2). We obtain

$$S = \sum_{p=1}^{M} \sum_{q=1}^{m} \left(\frac{m m_0}{\beta} \left(1 - \cos \frac{2\pi q}{m} \right) + \frac{m_0 \beta}{2m} \Omega_p^2 \right) |v_{p,q}|^2 + \sum_{p=1}^{M} \sum_{q=1}^{m} v_{p,q} \Lambda_p w_{-p,-q} , \tag{6.4a}$$

where

$$\Omega_p^2 \equiv M^{-1} \sum_{l=1}^{M} A_l \exp\left(\frac{2\pi i l p}{M} \right) \tag{6.4b}$$

and

$$\Lambda_p \equiv M^{-1} \sum_{l=1}^{M} B_l \exp\left(\frac{2\pi i l p}{M} \right) . \tag{6.4c}$$

It is now straightforward to integrate out the $v_{p,q}$ by changing variables and we find

$$Z_m = Z_m^{\mathrm{B}} Z_m^{\mathrm{Y}} , \tag{6.5}$$

where Z_m^{B} is the discrete path-integral approximant to the partition function of a free oscillator system and Z_m^{Y} is given by

$$Z_m^{\mathrm{Y}} = \sum_{\{y_{l,j}\}} \prod_{j,j'=1}^{m} \langle y_{1,j} \cdots y_{M,j} | \exp(-\beta H_1 / m) | y_{1,j+1} \cdots y_{M,j+1} \rangle \exp\left(\sum_{l,l'=1}^{M} F(j-j', l-l') \, y_{l,j} y_{l',j'} \right) \tag{6.6a}$$

and

$$F(j, l) = \frac{\beta^3}{4 m_0 m^4 M} \sum_{p=1}^{M} \sum_{q=1}^{m} \frac{\Lambda_p^2 \cos(2\pi l p / M) \cos(2\pi j q / m)}{1 - \cos(2\pi q / m) + \frac{1}{2} (\beta \Omega_p / m)^2} . \tag{6.6b}$$

As it is straightforward to evaluate Z_m^{B} numerically to very high accuracy (Schweizer et al., 1981; De Raedt and Lagendijk, 1983b) it is clear from (6.5) that the problem of calculating Z_m has been reduced to the problem of calculating Z_m^{Y}. The exponential factor in (6.6) represents the only effect of the coupling to the bosons on the system described by H_1 and can be interpreted as an effective interaction. In general the effective coupling $F(j, l)$ can be repulsive or attractive.

If $m \to \infty$ the sum over q in (6.6b) can be evaluated analytically and therefore it is a good approximation to replace (6.6b) by

$$F(j, l) = \frac{\beta^2}{4 m_0 m^2} \frac{1}{M} \sum_{p=1}^{M} \frac{\Lambda_p^2 \cos(2\pi p l / M)}{\Omega_p} \frac{\exp(-j \beta \Omega_p / m) + \exp(-(m-j) \beta \Omega_p / m)}{1 - \exp(-\beta \Omega_p)} , \tag{6.7}$$

if m is very large. If one starts from the continuum form of the path integral and eliminates the boson field, one obtains expression (6.7) directly (Feynman and Hibbs, 1965).

Since we are mainly interested in lattice models, for which a continuum path-integral does not make much sense, we will for reasons of internal consistency, work with the discrete form (6.6b) even though

the use of the continuum approximation (6.7) results in a more efficient computational scheme. The convergence properties of both the continuum and the discrete lattice path-integral of the harmonic oscillator have been compared (De Raedt and Lagendijk, 1983b) and it was found that the discrete lattice approach is superior from the point of view of accuracy.

6.2. Electron–phonon systems

6.2.1. Small polaron without phonon dispersion

We want to discuss a lattice model in which one fermion is coupled to a boson field, a polaron. In the theory of solids and liquids the polaron is a fundamental concept (Emin, 1982). The polaron we will treat in detail is described by the Holstein Hamiltonian, a lattice model originating from a tight-binding Hamiltonian. There have been many speculations about possible phase transitions ("self-trapping") in the ground state of continuum and lattice polarons. We will show that our results give some definite answers with respect to these long-standing problems. The polaron model we are considering is usually referred to as the Molecular Crystal Model (MCM) (Holstein, 1959b) and is also called the small-polaron Hamiltonian (Mahan, 1981). The major part of the physics of the small polaron is already displayed in the MCM without phonon dispersion.

We will first discuss this more simple MCM, that is to say without dispersion, for which numerical results can easily be obtained for all lattice dimensionalities. An analogous treatment of the molecular polaron, i.e. the two-site problem, can be found elsewhere (De Raedt and De Raedt, 1983, 1984). In the following section on more complicated polarons we will then deal shortly with the complication of phonon dispersion within the MCM. For simplicity we will now formulate the theory in one space dimension. The formulae for two-dimensional (2D) and three-dimensional (3D) systems can be derived by means of the same technique but we will not present them here. The model Hamiltonian reads

$$H = H_1 + H_2 + H_3, \tag{6.8a}$$

$$H_1 = -t \sum_{i=1}^{M} (c_i^+ c_{i+1} + c_{i+1}^+ c_i), \tag{6.8b}$$

$$H_2 = \frac{1}{2m_0} \sum_{i=1}^{M} p_i^2, \tag{6.8c}$$

$$H_3 = \frac{m_0 \Omega^2}{2} \sum_{i=1}^{M} x_i^2 + \lambda \sum_{i=1}^{M} x_i c_i^+ c_i, \tag{6.8d}$$

where Ω is the angular frequency of the Einstein oscillator, λ is the fermion–boson coupling strength, t is the kinetic energy associated with the nearest-neighbor hopping motion of the fermion, c_i^+ creates a fermion at site i, and c_i removes a fermion from site i. Hamiltonian (6.8) describes an electron coupled linearly to the phonon field of the site where the electron resides. This is the most extreme form of a short-range interaction. The phonons are dispersionless and consequently the only intersite communication is through the electron. Physical realizations of the d-dimensional model could be found in molecular crystals. One needs a molecular unit compatible with the lattice symmetry and having a non-degenerate internal mode, for instance, the breathing mode. As Hamiltonian (6.8) is of standard form (6.1) we can use the results obtained there. We need the following replacements,

$$\Omega_p \to \Omega\,, \tag{6.9a}$$

$$\Lambda_p \to \lambda\,, \tag{6.9b}$$

and if, as in chapter 4, the positions of the spinless fermions are denoted by $i_{\mu,j}$ we have

$$y_{l,j} = \sum_{\mu=1}^{N} \delta_{l,\, i_{\mu,j}}\,. \tag{6.9c}$$

Thus the matrix element in (6.6) is a special case ($N = 1$) of (4.10) and Z_m^{Y} becomes

$$Z_m^{\mathrm{Pol}} = \sum_{\{i_j\}} \rho(\{i_j\})\,, \tag{6.10a}$$

$$\rho(\{i_j\}) = \prod_{j,j'=1}^{m} I\Big(\frac{2\beta t}{m}, i_j - i_{j+1}\Big) \exp\{F(j-j', i_j - i_{j'})\}\,, \tag{6.10b}$$

where

$$F(j,l) = \delta_{l,0} \frac{\beta^3 \lambda^2}{4 m_0 m^4} \sum_{q=1}^{m} \frac{\cos(2\pi jq/m)}{1 - \cos(2\pi q/m) + \frac{1}{2}(\beta\Omega/m)^2}\,, \tag{6.10c}$$

and we omitted the irrelevant number-of-particles index μ from $i_{\mu,j}$. The calculation of the polaron contribution (6.10) is a non-trivial problem because the density function (6.10c) describes a peculiar 2D lattice system of m particles at the positions i_j interacting with each other. The first factor in (6.10b) represents an effective nearest-neighbor interaction, the second accounts for the retarded (in imaginary-time) long-range interactions caused by the fermion–boson coupling. In each row the real-space direction contains one electron interaction with the electrons in other rows. Note that for one particle it does not matter whether the particle is a boson or a fermion.

The partition function Z_m^{Pol} itself is not interesting, more relevant are the approximants to the energy, specific heat and derivatives of the polaron free energy $F_m^{\mathrm{Pol}} = -(1/\beta) \ln Z_m^{\mathrm{Pol}}$ with respect to the coupling λ. To simplify the notation a little bit, we will denote the expectation value of a quantity A taken in the ensemble defined by the density function $\rho(\{i_j\})$ by

$$\langle\langle A \rangle\rangle_m \equiv \frac{1}{Z_m^{\mathrm{Pol}}} \sum_{\{i_j\}} \rho(\{i_j\})\, A(\{i_j\})\,. \tag{6.11}$$

We have for the polaron energy (De Raedt and Lagendijk, 1983b)

$$E_m^{\mathrm{Pol}} = K_m + V_m\,, \tag{6.12a}$$

$$K_m = -\frac{t}{m} \sum_{j=1}^{m} \Big\langle\Big\langle \frac{I(2\beta t/m, i_j - i_{j+1} - 1) + I(2\beta t/m, i_j - i_{j+i} + 1)}{I(2\beta t/m, i_j - i_{j+1})} \Big\rangle\Big\rangle_m\,, \tag{6.12b}$$

$$V_m = -\sum_{j,j'=1}^{m} \Big\langle\Big\langle \frac{\partial F(j-j', i_j - i_{j'})}{\partial \beta} \Big\rangle\Big\rangle_m\,. \tag{6.12c}$$

If a model is expected to show non-analytic behavior in some of its thermodynamic functions or in its ground-state properties one would like to know the character of these non-analyticities. One can think of first-order type behavior with coexistence of different phases or of second-order type behavior with critical fluctuations or of even different (e.g. higher-order) transitions. In such a case one investigates the susceptibilities of the driving fields. Since the electron–phonon coupling is the driving force in the occurrence of self-trapping, it is obvious that one should focus on the derivatives of the polaron free energy F_m^{Pol} with respect to the coupling parameter. If there are no peculiarities in these quantities there is no transition. The first derivative $\partial F_m^{\mathrm{Pol}}/\partial\lambda$ is equal to the coupling energy and the second derivative $\partial^2 F_m^{\mathrm{Pol}}/\partial\lambda^2$ is related to the coupling-energy–coupling-energy Kubo susceptibility. Straightforward algebra yields

$$\frac{\partial F_m^{\mathrm{Pol}}}{\partial\lambda} = -\frac{2}{\beta\lambda}\sum_{j,j'=1}^{m}\langle\langle F(j-j', i_j - i_{j'})\rangle\rangle_m \tag{6.13a}$$

and

$$\frac{\partial^2 F_m^{\mathrm{Pol}}}{\partial\lambda^2} = -\frac{4}{\beta\lambda^2}\sum_{\substack{j,j'=1\\k,k'=1}}^{m}\langle\langle F(j-j', i_j - i_{j'})\,F(k-k', i_k - i_{k'})\rangle\rangle_m - \frac{1}{\lambda}\frac{\partial F_m^{\mathrm{Pol}}}{\partial\lambda} + \beta\left(\frac{\partial F_m^{\mathrm{Pol}}}{\partial\lambda}\right)^2. \tag{6.13b}$$

One contribution to $\partial^2 F_m^{\mathrm{Pol}}/\partial\lambda^2$ is proportional to $\partial F_m^{\mathrm{Pol}}/\partial\lambda$ and subtracting this background defines the susceptibility χ_m^{Pol},

$$\chi_m^{\mathrm{Pol}} = -\frac{\partial^2 P_m^{\mathrm{Pol}}}{\partial\lambda^2} - \frac{1}{\lambda}\frac{\partial F_{\mathrm{m}}^{\mathrm{Pol}}}{\partial\lambda}. \tag{6.14}$$

In order to gain additional insight in the physics of the polaron we also calculate the electron–phonon correlation functions

$$\hat{C}(l) = \sum_{n=1}^{M}\frac{\mathrm{Tr}\,\mathrm{e}^{-\beta H}c_n^{+}c_n x_{n+l}}{\mathrm{Tr}\,\mathrm{e}^{-\beta H}}. \tag{6.15a}$$

Analytic elimination of the phonons yields

$$\hat{C}(l) = \lim_{m\to\infty}\hat{C}_m(l), \tag{6.15b}$$

$$\hat{C}_m(l) \equiv \frac{2}{\beta\lambda}\sum_{j'=1}^{m}\langle\langle F(j-j', i_j - i_{j'} + l)\rangle\rangle_m. \tag{6.15c}$$

The formalism presented in this subsection so far, giving rise to highly non-trivial actions, is of little use unless we can calculate the various observables for different model parameters. Although in a strict sense the density function (6.10b) is not a density function of a classical model, one can still use standard Monte Carlo procedures, as outlined in chapter 2 to estimate the expectation values within a certain

statistical accuracy. As matter of fact, compared with the Monte Carlo simulations of the systems treated in chapters 4 and 5, simulating (6.10) is really a simple excercise.

To find out what typical values of m are needed we first studied the convergence of the boson contribution to the partition function Z_m^{B} as a function of m. For the dispersionless model Z_m^{B} is given by

$$Z_m^{\mathrm{B}} = \left[\prod_{q=1}^{m}\left(1-\cos\frac{2\pi q}{m}+\frac{\beta^2\Omega^2}{2m^2}\right)^{-1/2}\right]^M. \tag{6.16}$$

In table 6.1 we show some results for this boson system for a typical set of model parameters. Note the slow convergence of the specific heat. From these studies we can deduce a minimum value of m such that for each inverse temperature m the exact results of the boson system are reproduced within a specified error. Each and every Monte Carlo algorithm should be tested as much as possible. In this particular problem analytic (or numerically rigorous) results can be derived for weak coupling and for strong coupling. Other important tests are the reproduction of the (trivial) results for non-interacting systems of different sizes. Note that in our approach the exact results for the non-interacting system will be generated for any value of m. In our final simulations we chose $L = 32$ (variation of the linear dimension only has a very small effect on the results). The Monte Carlo calculation was very efficient: a typical run ($d = 1$, $\beta = 5$, $m = 32$, $t = 1$, $\lambda = 3$, and 50000 samples) took 40 minutes CPU time on a Digital Equipment Corporation VAX 11/780.

We will only present a few results. In fig. 6.1 we show the energy and the free-energy derivatives $\partial F_m^{\mathrm{Pol}}/\partial\lambda$ and $\partial^2 F_m^{\mathrm{Pol}}/\partial\lambda^2$ as a function of λ, keeping β constant. The value of β is chosen so high that the system is almost in the ground state. In the same figure the outcome of weak- and strong-coupling theories is presented. In fig. 6.2 the susceptibility χ_m^{Pol} is displayed. The increased fluctuations at low temperatures for $2 < \lambda < 3$ are indicative of a possible transition in the ground state. Qualitatively the same results were found for 2D and 3D. A clear picture of what is going on can be obtained by inspecting fig. 6.3. In this figure the normalized correlation function $C_m(\boldsymbol{l}) \equiv \hat{C}_m(\boldsymbol{l})/\hat{C}_m(0)$ is displayed for $|\boldsymbol{l}| = 1$. The vector notation is introduced because we now explicitly consider results for all

Table 6.1
The free energy per site F_m^{B}, energy per site E_m^{B} and specific heat per site C_m^{B} of the free boson system (i.e. the harmonic oscillator) as a function of m. The inverse temperature $\beta = 5$ and the oscillator frequency $\omega = 1$. The rigorous results correspond to $m = \infty$. For small m the free energy is larger than the thermal energy

m	$F_m^{(1)}$	$E_m^{(1)}$	$C_m^{(1)}$
1	0.3219	0.2000	1.0000
2	0.4160	0.3220	1.1338
3	0.4530	0.3923	0.9630
4	0.4703	0.4316	0.7692
8	0.4907	0.4842	0.3838
16	0.4966	0.5008	0.2298
32	0.4981	0.5053	0.1859
64	0.4985	0.5064	0.1746
128	0.4986	0.5067	0.1717
∞	0.4986	0.5068	0.1707

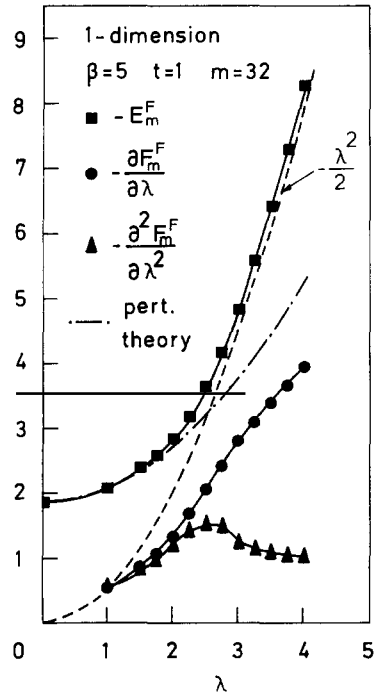

Fig. 6.1. The energy, the first derivative of the free energy with respect to λ and the second derivative of the free energy with respect to λ as a function of the coupling λ of the small polaron in 1D. Also shown are the results of weak- and strong-coupling theories.

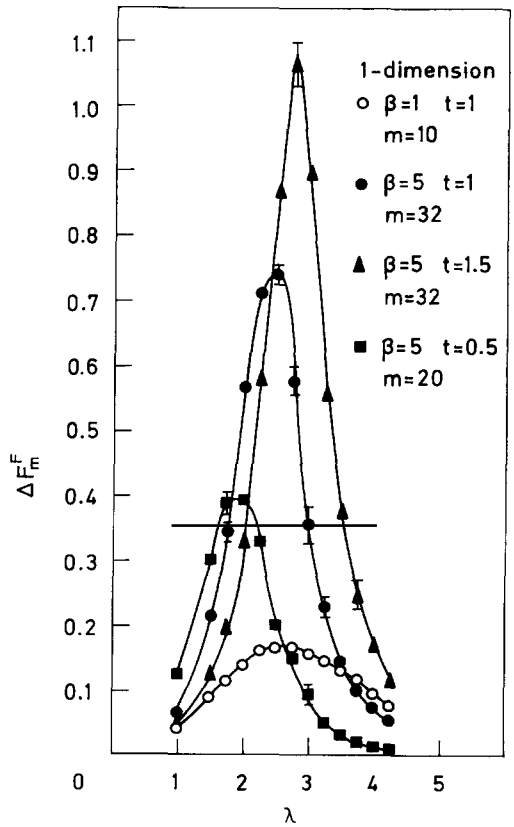

Fig. 6.2. The coupling dependence of the fluctuation χ_m^{Pol} for several temperatures and hopping energies t of the small polaron in 1D.

dimensionalities. The sharp drop of this correlation function beyond a certain value of λ to an almost vanishing value means that the polaron's spatial extension is essentially over only one site, the polaron is self-trapped. Basically we find the same features for all dimensionalities: beyond a "critical" value of λ self-trapping occurs. It seems that this self-trapping transition is a continuous transition with enhanced fluctuations. Whether or not derivatives of the free energy show really singular behavior cannot be ascertained from Monte Carlo data alone. An accurate estimate of the observed critical coupling constant can be obtained by equating the weak- and strong-coupling expressions for the ground-state energy and solving this equation for λ. Our results are in disagreement with a variational calculation that predicts a first-order transition in 3D (Emin, 1973). An interesting region from a physical point of view in parameter space of the MCM is the adiabatic regime. In our formulations the adiabatic model is obtained by putting $F(j, l) = \beta\lambda^2\delta_{l,0}/2m_0\Omega^2m^2$. Simulations in the adiabatic regime did not give significantly different results. In the adiabatic limit, the two-site model (the molecular polaron), the 1D MCM, 2D MCM and 3D MCM exhibit very similar properties as far as the transition from the weak-coupling to the strong-coupling regime is concerned. For more details we refer to (De Raedt and Lagendijk, 1981a, 1983b, 1984; De Raedt and De Raedt, 1983, 1984).

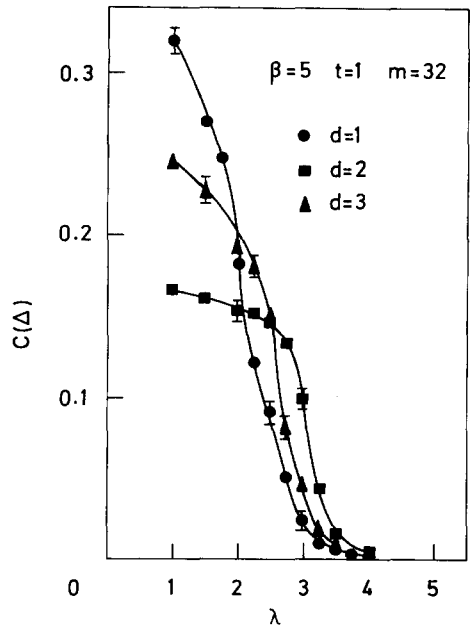

Fig. 6.3. The normalized nearest-neighbor fermion–boson correlation functions for 1-, 2- and 3-dimensional polaron motion as a function of the coupling λ. Δ is a unit vector of the d-dimensional lattice.

6.2.2. More complicated polarons

The first sophistication which comes to mind is the introduction of phonon dispersion in the MCM. We will go in some detail for the 1D model. The phonon dispersion in the MCM is introduced by allowing for nearest-neighbor coupling of the oscillators. The coupling Hamiltonian can now be written as

$$H_3 = \frac{m_0\omega_0^2}{2}\sum_{i=1}^{M} x_i^2 + \frac{m_0\omega_1^2}{2}\sum_{i=1}^{M} x_i x_{i+1} + \lambda \sum_{i=1}^{M} x_i c_i^+ c_i \,. \tag{6.17}$$

The total Hamiltonian, $H_1 + H_2 + H_3$ is still of standard form (6.1) and the standard result (6.4) can be carried over. We need the following replacements

$$\Omega_p \to \omega_0^2 + \omega_1^2 \cos(2\pi p/M)\,, \tag{6.18a}$$

$$\Lambda_p \to \lambda\,. \tag{6.18b}$$

Using the general result (6.6) for the partition function one easily arrives at the conclusion that we only have to substitute

$$F(j,l) = \frac{\beta^3\lambda^2}{4m_0m^4M}\sum_{p=1}^{M}\sum_{q=1}^{m} \frac{\cos(2\pi lp/M)\cos(2\pi jq/m)}{1-\cos(2\pi q/m)+\frac{1}{2}(\beta\Omega_p/m)^2}\,, \tag{6.19}$$

instead of (6.10c) into (6.10b). Note the complication that has developed into expression (6.19), viz. the

new function $F(j, l)$ contains a summation over a $(d+1)$-dimensional lattice to be compared with a one-dimensional (in the Trotter direction) sum in case of the dispersionless MCM (see (6.10c)). As it is very inefficient to calculate $F(j, l)$ during the simulation itself, the standard manner to deal with this problem is to calculate *all* $F(j, l)$ and to store them. For a large 3D model, the computation time and storage requirements for $F(j, l)$ make the calculation of $F(j, l)$ more expensive than the simulation itself unless a simple parametrization of the phonon spectrum is possible. Remark that in the adiabatic limit one has $F(j, l) = (\beta\lambda^2/2m_0m^2M)\sum_{p=1}^{M}\Omega_p^{-2}\cos(2\pi lp/M)$, so that things are somewhat simpler.

The phonon dispersion we have introduced is best described by introducing the gap parameter g, $g^2 = 1 - \omega_1^2/\omega_0^2$ and by taking $\Omega_{p=0} = 1$. The dispersionless MCM is recovered by setting $g = 1$. We have performed extensive simulations of the 1D MCM with various gaps. At very low temperatures we again find strongly enhanced fluctuations and a steep change of correlation functions measuring the spatial size of the polaron. The self-trapping process is more difficult to visualize since the phonon dispersion destroys the neat picture of a small polaron trapped on one site only. In fig. 6.4 the susceptibility χ_m^{Pol} is displayed, and in fig 6.5 the critical value of the coupling constant, $\lambda_c(g)$, is shown as a function of the gap g. Apparently $\lambda_c(g) \to 0$ when $g \to 0$.

One should be very careful by interpreting the data in terms of a localization transition. The transition from a band-like motion for small λ to a self-trapped state for large λ is sometimes referred to as a localization transition. One should realize that this does not imply that the translational symmetry is broken. From our simulation data, we conclude that there is no evidence for this to occur (De Raedt and Lagendijk, 1984). All the correlation functions we have introduced and which we have measured with the Monte Carlo method are manifestly translationally invariant.

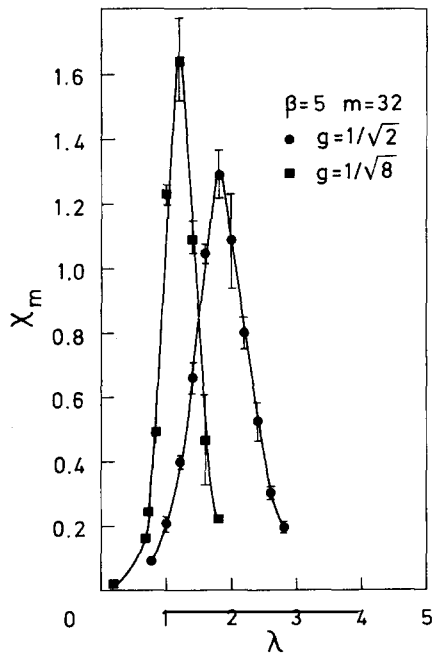

Fig. 6.4. The electron–phonon coupling-energy susceptibility χ_m^{Pol} as a function of the electron–phonon coupling λ of the small polaron in 1D with phonon dispersion. A discontinuity in χ_m^{Pol} would indicate that the polaron undergoes a second-order ground-state phase transition as the electron–phonon interaction λ increases from zero to infinity.

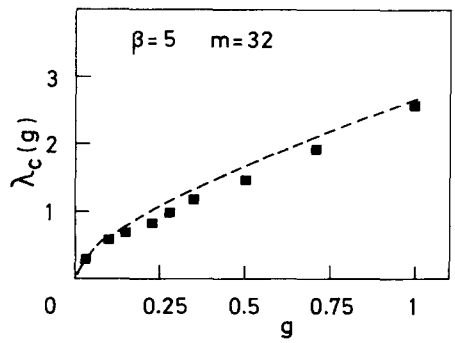

Fig. 6.5. The critical value $\lambda_m(g)$ of the electron–phonon coupling as a function of the phonon gap g at the Brillouin zone boundary. The dashed line is a theoretical result obtained from comparison of weak and strong coupling expansions for the small polaron Hamiltonian.

The most complicated class of polarons that still can be handled with the phonon-elimination trick are those polarons which have dispersion in both the phonon energy and the coupling energy. Furthermore a polaron can also be defined in continuum space rather than on a discrete lattice. Shoji and Tokuda (Shoji and Tokuda, 1981) have given a handy compilation of some popular continuum polarons. The generic Hamiltonian is given by

$$H = \frac{p^2}{2m_0} + \sum_{\boldsymbol{k}} Q(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})(a^+_{-\boldsymbol{k}} + a_{\boldsymbol{k}}) + \sum_{\boldsymbol{k}} \omega(\boldsymbol{k})\, a^+_{\boldsymbol{k}} a_{\boldsymbol{k}} . \tag{6.20}$$

In table 6.2 the values of $Q(\boldsymbol{k})$ and $\omega(\boldsymbol{k})$ are presented for four types of polarons. The optical polaron in 3D is the Fröhlich polaron (Appel, 1968), to which Feynman applied his well-known variational principle (Feynman, 1955; Feynman and Hibbs, 1965; Feynman, 1972). Another problem which can be cast in the standard form (6.20) in 2D is the interaction of an electron on a helium (^{4}He) film with the ripplons (Jackson and Platzman, 1981). A very similar problem is the coupling of a hydrogen atom to a helium (^{4}He) film (Wilson and Kumar, 1983).

Continuum models usually have to be regularized by introducing appropriate cut-offs. The Fröhlich Hamiltonian is cut-off free in 3D. All these models can in principle be treated in the same way like we did for the MCM. Continuum models have a Gaussian rather than the modified Bessel function as the propagator of the free electron. As we indicated in subsection 6.2.1 in the continuum approximation the function representing the retarded interaction, $F(j, l)$, can be simplified considerably because one summation in the Trotter direction can be done analytically. A nice feature of isotropic continuum models is that evaluation of d-dimensional integrals reduces to doing a one-dimensional integral such that there are no practical problems to calculate and store $F(j, l)$, even in 3D.

Many of the continuum models are expected to show phase transitions in the ground state. Continuum models will in many cases exhibit quite different behavior when compared with their lattice counterparts. In contrast to lattice models with band-type motion, in continuum space kinetic energy operators are not bounded from above. This has important consequences for self-trapping effects. In relation to self-trapping effects, continuum models and continuum versions of the MCM have been extensively studied by means of various approximate theories such as the adiabatic approximation (Emin and Holstein, 1976), and the Feynman variational approach (Sumi and Toyozawa, 1973; Luttinger and Lu, 1980). It is difficult to assess the value of simple mean-field theories which usually predict phase transitions, but as was shown by Peeters and Devreese (Peeters and Devreese, 1982) for the Fröhlich polaron, compare unfavorably with the best variational calculation (Feynman, 1955).

Table 6.2
Different polaron types. The interaction parameters refer to (6.20), s is the speed of sound and the α_i's are dimensionless coupling parameters (after Shoji and Tokuda, 1981)

Type of polaron	$Q(k)$	$\omega(k)$	units
Optical	$(2\sqrt{2}\pi\alpha_1)^{1/2}k^{-1}$	1	$m_0 = \hbar = \omega = 1$
Piezo-electric	$(2\sqrt{2}\pi\alpha_2)^{1/2}k^{-1/2}$	k	$m_0 = \hbar = s = 1$
Acoustic by the deformation potential	$(2\sqrt{2}\pi\alpha_3)^{1/2}k^{1/2}$	k	$m_0 = \hbar = s = 1$
Optical by the deformation potential	$(2\sqrt{2}\pi\alpha_4)^{1/2}$	1	$m_0 = \hbar = \omega = 1$

We have shown that a wealth of information can be obtained by combining the analytic technique of integrating out the phonons with the numerical technique of Monte Carlo simulation. We think that this combined method should be applied as much as possible to these important models. The reader who has gone through the detailed results on the MCM should be able to apply this technique to any model of the standard form. We have to give a warning here, since sometimes the results might be disappointing as we, and others (Becker, Gerlach and Schliffke, 1983) found out for the Fröhlich polaron in the strong-coupling regime. The function $F(j, l)$ might be of singular nature and although the model might still be well-defined, this could make the application of the Monte Carlo technique impossible.

In our opinion it would appear that the best break-up of polaron Hamiltonians of standard type (6.1) or (6.20) is the one that allows for analytic elimination of the phonons. However, in physics many polarons exist for which the analytic elimination is either more difficult or completely impossible. An example of the first class is a polaron where the coupling with the lattice vibrations is quadratic in the vibration amplitude. This would apply to condensed-matter systems where the linear term vanishes because of symmetry or where the linear term is small. There is a long-standing suggestion that the quadratic coupling to librons in TTF-TCNQ type compounds is very important in understanding the transport properties (Gutfreund and Weger, 1977). The model proposed by Gutfreund and Weger is however more complicated, and we do not know of a situation where a simple quadratic coupling to the electron density is relevant. Complications of much more severity are showing up when the electron–phonon coupling is more complex than a simple coupling to the electron density. This is of course possible because we work in a reduced Hilbert space, and parameters in a second-quantized Hamiltonian are averaged quantities of the full Hilbert space. In this way we can get peculiar couplings in a second-quantized Hamiltonian. A very fashionable polaron Hamiltonian of this kind, is the polaron in which the hopping integral t is modulated by the phonons. This model is now usually referred to as the SSH (Su, Schrieffer and Heeger, 1980) model, although the model was introduced long before SSH applied it to polyacetylene (Holstein, 1959a). The 1D SSH Hamiltonian is given by

$$H = H_1 + H_2 + H_3\,, \tag{6.21a}$$

$$H_1 = -t \sum_{i=1}^{M} (c_i^+ c_{i+1} + c_{i+1}^+ c_i)\,, \tag{6.21b}$$

$$H_2 = \alpha \sum_{i=1}^{M} (c_i^+ c_{i+1} + c_{i+1}^+ c_i)(x_{i+1} - x_i)\,, \tag{6.21c}$$

$$H_3 = \frac{m_0 \omega_0^2}{2} \sum_{i=1}^{M} x_i^2 + \frac{m_0 \omega_1^2}{2} \sum_{i=1}^{M} x_i x_{i+1} + \lambda \sum_{i=1}^{M} x_i c_i^+ c_i\,. \tag{6.21d}$$

The restriction to linear modulation of t is mainly caused by the necessity to keep the model as simple as possible. One should be aware of the fact that H_2 is the result of a Taylor expansion and is only correct for small vibrational amplitudes. This is very important for functional integration techniques, because in these methods one integrates over all amplitudes. To circumvent this problem one could use some empirical patch by writing the transfer integral $t_{i,j}$ between site i and j as

$$t_{i,j} = t_0 \exp\{-\kappa |x_i - x_j + (i-j)a|\}\,, \tag{6.22}$$

where t_0 and κ should be fitted to the overlap integrals, and a is the lattice constant. For instance the transfer integral between neighboring sites would then be $t = t_0 e^{-\kappa a}$. Hirsch and Fradkin (Hirsch and Fradkin, 1983) used a more ad hoc procedure of setting the local two-site hopping energy $t - \alpha(x_{i+1} - x_i)$ equal to zero whenever $\alpha(x_{i+1} - x_i) < t$. For the SSH model it is not possible to eliminate the phonon degrees of freedom analytically. This means that the very nice separation of fermion and phonon degrees of freedom is not possible any longer. This is very bad from a Monte Carlo point of view because if we would be able to simulate this model it was very difficult to extract the polaron information out of these computer experiments since it would be drowned in the noise of the many-oscillator degrees of freedom. This is particularly true for additive quantities like energy and specific heat where the polaron contribution, which is of $\mathcal{O}(1)$, is defined as the difference of two contributions of $\mathcal{O}(M)$. Since the Feynman trick does not work for the SSH model one must think of other ways of obtaining a path-summation representation. An obvious possibility is a real-space break-up, either a simple one like suggested by Suzuki (Suzuki, 1976a) or the more ingenious CBD introduced by Barma and Shastry (Barma and Shastry, 1977, 1978). We have treated both decompositions in detail in chapter 4 for fermions and in chapter 5 for spins. Hirsch and Fradkin (Hirsch and Fradkin, 1982; Fradkin and Hirsch, 1983) used the Barma and Shastry break-up in relation with the SSH model, but their work refers to a many-electron situation and will be discussed in the next section. For a many-electron problem the elimination of the oscillator degrees of freedom is not a prerequisite for a sensible Monte Carlo approach.

6.2.3. Many-polaron systems

In many situations the single-polaron case is not a satisfactory description of the real situation. The interaction between the polarons should not be neglected, and ought to be taken into account right from the start. One should realize that many-body electron–phonon models display a rich variety of exotic behavior including superconductivity, negative U-centers, exciton condensation etc. A general Monte Carlo strategy for all these models is of course not possible. To get an impression of what models could eventually be simulated is much easier. We visualize an interacting many-body system coupled to a harmonic-oscillator bath. If the interacting many-body system without coupling cannot be simulated with Monte Carlo methods, this certainly holds for the total, more complex, system. With the information given in this review up to now this implies that many-fermion systems in $d > 1$ will be very difficult, if not impossible for other than half-filled band cases. As less ambitious scheme would be to probe few-body polaronic models like bipolarons. There is a lot of fundamental physics connected with these entities, and Monte Carlo simulations are quite feasible, including few-fermion problems in all dimensions. A complication in the many-polaron problem is the occurrence of spin. We will only consider models for which the spin-dependent interactions are either absent or quite simple like in the Hubbard model. If the many-body polaron Hamiltonian is of standard form (6.1) the analytic elimination of the phonons is still possible, and the result is still given by (6.6). First we deal with a spinless 1D N-fermion system with nearest-neighbor hopping kinetic energy only. Combination of (4.10) and (6.6) gives

$$Z_m^{\mathrm{Pol}} = \sum_{\{i_{1,j} < \cdots < i_{N,j}\}} \sum_{\{P_j\}} \left[\prod_{j=1}^{m} \prod_{\mu=1}^{N} \mathrm{sign}(P_j)\, I\left(\frac{2\beta t}{m}, i_{\mu,j} - i_{P_j\mu,j+1}\right) \right] \exp\left(\sum_{j,j'=1}^{m} \sum_{\mu,\nu=1}^{N} F(j-j', i_{\mu,j} - i_{\nu,j'}) \right), \tag{6.23}$$

where the general expression of $F(j, l)$ is again given by (6.6b). It is no problem to incorporate the effect

of a direct interaction (short range or long range) between the electrons, and (4.13) gives an example of how to include nearest-neighbor repulsion. A considerable simplification of (6.23) is possible if we limit ourselves to the MCM (without dispersion), as should be clear now from the calculations presented in subsection 6.2.2. The principal problem of (6.23) is of course the appearance of negative contributions due to the sign of the permutations. It is illustrative to write out explicitly the sum over permutations for the case of two fermions

$$Z_m^{\mathrm{Pol}} = \sum_{\{i_{1,j}<i_{2,j}\}} \Bigg(\prod_{j=1}^{m} \prod_{\mu=1}^{N} \Big[I\Big(\frac{2\beta t}{m}, i_{1,j} - i_{1,j+1}\Big) I\Big(\frac{2\beta t}{m}, i_{2,j} - i_{2,j+1}\Big) - I\Big(\frac{2\beta t}{m}, i_{1,j} - i_{2,j+1}\Big) \times I\Big(\frac{2\beta t}{m}, i_{2,j} - i_{1,j+1}\Big) \Big] \Bigg) \exp\Big(\sum_{j,j'=1}^{m} \sum_{\mu,\nu=1}^{N} F(j-j', i_{\mu,j} - i_{\nu,j'}) \Big) . \tag{6.24}$$

In (6.24) the possibility of negative signs has been made explicit by writing out the fermion determinant. For an infinite 1D system there are no negative contributions because $I(z, l) = I_l(z)$, $I_l(z) > I_{l+1}(z)$ if $z > 0$ (De Sitter and Goovaerts, 1973) and the sum in (6.24) is over an ordered set. For two particles we expect the sign problem to be mild even if $d > 1$ (see also chapter 7). For instance one could sample in the ensemble defined by the absolute value of the determinant. This is a better procedure than to sample in an ensemble defined by the permanent (Takahashi and Imada, 1983). The latter procedure was used by us for a many-fermion system as explained in chapter 4. We were forced to do so because calculation of large determinants becomes very time consuming. Although up to now virtually no Monte Carlo results have been published on few-fermion polaron problems we see no reason why this cannot be done. There would be no problems at all for few-boson systems. For the real many-fermion problem in 1D our method used for the fermion-only Hamiltonians can be used, although it has not been done up to now. Introduction of spin is simple if one uses the simple but clever trick of the Santa Barbara group (Hirsch et al., 1982) to work in an ensemble defined by constant number of up and constant number of down spins. It is absolutely necessary in this case, however, to know whether or not this sector contains the ground state. We have explained this procedure in chapter 4, and the integration of this recipe in the many-polaron action is straightforward.

If the phonons cannot be eliminated, or if one prefers not to do so, one can use a real-space break-up for the many-polaron Hamiltonian. As was toughed upon earlier this is a little tricky for few-polaron situations, since phonons overwhelmingly dominate the situation, but becomes quite feasible in 1D if the density of the fermions becomes of $\mathcal{O}(1)$, rather than being of $\mathcal{O}(1/M)$. A difficulty one then encounters is the difference in energy scale between electron and phonon motion.

Hirsch and Fradkin have used the Barma-and-Shastry checkerboard break-up to study some 1D many-polaron models. The models they have investigated are the 1D dispersionless MCM and the 1D SSH models both for electron density $\rho = \frac{1}{2}$. The two systems were studied with and without spin degrees of freedom. Introduction of the oscillator degrees of freedom makes the simulation much more involved since in the checkerboard approach, they have to be simulated also. Nevertheless these models are important and any new information on them is of course welcome.

In the checkerboard approach, the approximant to the partition function of the MCM is given by

$$Z_m = c\Big(\frac{\beta}{m}\Big)^{-mM/2} \mathrm{Tr} \int \Big[\prod_{i=1}^{M} \prod_{j=1}^{m} \mathrm{d}x_{i,j} \Big] \exp[-S(\{x_{i,j}\})] \exp\Big[-\frac{\beta\lambda}{m} \sum_{i=1}^{M} \sum_{j=1}^{m} \sum_{\sigma} x_{i,j}(n_{i,\sigma} - \tfrac{1}{2}) \Big] \times$$

$$\times \exp\left[\frac{\beta t}{m}\sum_{l=1}^{M/2}\sum_{j=1}^{m}\sum_{\sigma}(c^{+}_{2l,\sigma}c_{2l+1,\sigma}+c^{+}_{2l+1,\sigma}c_{2l,\sigma})\right]$$

$$\times \exp\left[\frac{\beta t}{m}\sum_{l=1}^{M/2}\sum_{j=1}^{m}\sum_{\sigma}(c^{+}_{2l+1,\sigma}c_{2l+2,\sigma}+c^{+}_{2l+2,\sigma}c_{2l+1,\sigma})\right], \tag{6.25a}$$

where

$$S(\{x_{i,j}\})=\frac{mm_0}{2\beta}\sum_{i=1}^{M}\sum_{j=1}^{m}(x_{i,j}-x_{i,j+1})^2+\frac{m_0\beta\Omega}{2m}\sum_{i}^{M}\sum_{j=1}^{m}x_{i,j}^2, \tag{6.25b}$$

c is an unimportant constant and the sum over σ runs over the spin states. A thorough discussion of the physics of these complicated polaron models is far beyond the scope of this review. We refer to the extensive original work by Hirsch and Fradkin (Hirsch and Fradkin, 1982, 1983; Fradkin and Hirsch, 1983) in which more than Monte Carlo results are being discussed. Hirsch and Fradkin are interested in the occurrence of the Peierls distortion. To get insight in this phenomenon they sample among other things the (staggered) phonon order parameter defined by $m_p=(1/M)\Sigma_{i=1}^{M}\langle x_i\rangle$. A non-zero value of this order parameter indicates a ground state with broken symmetry, a dimerized state. Hirsch and Fradkin conclude that in the spinless case there is a finite value for the coupling constant at which the ground state gets dimerized. This seems to be different for the spin-carrying model because Hirsch and Fradkin conclude from their simulation data that the critical value of coupling constant is zero. For more details on these polaron models and for a discussion of the influence of the Coulomb (on-site and nearest-neighbor) interaction we refer the reader to the original papers (Hirsch and Fradkin, 1982, 1983; Fradkin and Hirsch, 1983; Hirsch, 1983a).

7. Fermions and bosons in 2 and 3 dimensions

In this chapter we consider the problem of calculating the partition function for a fixed number (N) of free fermions or bosons. Thus we want to calculate the properties of free fermions and bosons in the canonical ensemble. Usually one allows for fluctuations in the number of particles because then it is possible to find a closed-form expression for the (grand-canonical) partition functions. The average number of particles in the system can be set to the required number by making a suitable choice for the chemical potential. In this way all interesting physical properties of large systems of free fermions or bosons have been obtained (Landau and Lifshitz, 1980; Huang, 1963). Nevertheless it is still worthwhile to perform the more difficult calculations in the canonical ensemble. Not only do these calculations yield a lot of (numerically exact) reference data that should be reproduced by any other computational method but they also give useful information about the N-dependence of physical properties. Furthermore it allows us to study quantitatively the "minus-sign" problem that seems to be inevitable when treating 2D and 3D fermion systems. Since these aspects can be studied by numerically exact calculations for the case of free particles only, this chapter is mainly of pedagogical interest.

We first discuss the lattice model. To simplify the notation, we derive the expressions for the one-dimensional case but we give final expressions that are valid for the d-dimensional system. In order to compare lattice- and box-models it is expedient to calibrate units of energy in both models. Since

$$H_{\text{Lattice}} = -2t \sum_k \cos k c_k^+ c_k \approx -2tN + t \sum_k k^2 c_k^+ c_k + \cdots \qquad \text{and} \qquad H_{\text{Box}} = (2m_0)^{-1} \sum_k k^2 c_k^+ c_k$$

we will put $(2m_0)^{-1} = t$ in the rest of this chapter. As was shown in chapter 4 the partition function of N free fermions on a chain of M is given by

$$Z_N^{\text{Lattice}} = \frac{1}{N!} \sum_P \sum_{\{y_l\}} \text{sign}(P) \prod_{l=1}^{N} I(2\beta t, y_l - y_{Pl}) . \tag{7.1}$$

For any given permutation P the sum over all positions $y_l = 1, \ldots, M$ can be carried out analytically. To show this we choose one of the $N!$ permutations of the N particles and call it $\bar{P}$. Each subset of the set $\{1, \ldots, N\}$ that is invariant under the repeated action of the permutation $\bar{P}$ is called a cycle, the number of elements in the cycles is called the length of the cycle. For example, the permutation

$$\bar{P} \equiv \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 3 & 4 & 6 \end{pmatrix}, \tag{7.2}$$

can be represented by the cycles (12)(354)(6). In general a permutation consists of C_l cycles of length l where

$$\sum_{l=1}^{N} l C_l = N; \qquad 0 \le C_l \le N. \tag{7.3}$$

The number of different permutations that can be constructed from C_l cycles of length l is given by (Abramowitz and Stegun, 1965; Feynman, 1972)

$$M(\{C_l\}) = N! \left(\prod_{l=1}^{N} C_l!\, l^{C_l} \right)^{-1}. \tag{7.4}$$

For any P represented by the set $\{C_l ; l = 1, \ldots, N\}$,

$$Z(P) \equiv \sum_{\{y_l\}} I(2\beta t, y_l - y_{Pl})$$

is exactly the same. Taking the permutation $\bar{P}$ as an example we find

$$Z(\bar{P}) = I(2\beta t, 0) \sum_{\{y_l\}} I^2(2\beta t, y_1 - y_2)\, I(2\beta t, y_3 - y_5)\, I(2\beta t, y_5 - y_4)\, I(2\beta t, y_4 - y_3) . \tag{7.5}$$

Taking for instance $\tilde{P} = (123)(46)(5)$ we get

$$Z(\tilde{P}) = I(2\beta t, 0) \sum_{\{y_l\}} I^2(2\beta t, y_4 - y_6)\, I(2\beta t, y_1 - y_2)\, I(2\beta t, y_2 - y_3)\, I(2\beta, y_3 - y_1) , \tag{7.6}$$

which is exactly the same as (7.5) if we change variables. As illustrated by this simple example the contribution of a cycle of length l is given by

$$\mathscr{L}_l = \sum_{y_1=1}^{M} \cdots \sum_{y_l=1}^{M} I(2\beta t, y_1 - y_2) \cdots I(2\beta t, y_{l-1} - y_l)\, I(2\beta t, y_l - y_1)\,, \tag{7.7}$$

which is readily evaluated by repeated use of the summation formula

$$\sum_{y=1}^{M} I(a, x-y)\, I(b, y-z) = I(a+b, x-z)\,, \tag{7.8}$$

and yields

$$\mathscr{L}_l = M I(2\beta l t, 0)\,. \tag{7.9}$$

Since only cycles of even length change the parity of the permutation the final result for the partition function of a d-dimensional fermion lattice model can be written as

$$Z_N^{\text{Lattice}} = \frac{1}{N!} \sum_{\{C_l\}} \prod_{l=1}^{N} (-1)^{(l+1)C_l} \mathscr{L}_l^{dC_l} M(\{C_l\})\,, \tag{7.10}$$

where the sum is over all possible solutions of (7.3). As usual the expressions for the boson partition function is obtained by removing the minus sign in (7.10). The main advantage of summing over cycles instead of permutations is that for $N \geq 4$ the number of contributions to Z_N^{Lattice} has been reduced considerably. Indeed it can be shown that for large N the number of solutions of (7.3) is approximately given by $(4\sqrt{3}N)^{-1} \exp(\pi\sqrt{2N/3})$ whereas a brute force calculation of Z_N^{Lattice} requires the evaluation of

$N! \approx \sqrt{2\pi} N^{N+1/2} \exp(-N + 1/12N)$ terms (Abramowitz and Stegun, 1965).

We now turn to the problem of fermions (bosons) in a box, subject to periodic boundary conditions. For particles moving on a ring of length L we have

$$Z_N^{\text{Box}} = \frac{1}{N!} \sum_{P} \sum_{\{n_l\}} \text{sign}(P) \prod_{j=1}^{N} \delta_{n_j, n_{P_j}} \exp\left(-\frac{4\pi^2 \beta l t n_j^2}{L^2}\right). \tag{7.11}$$

Generalizing to d dimensions gives

$$Z_N^{\text{Box}} = \frac{1}{N!} \sum_{\{C_l\}} \prod_{l=1}^{N} (-1)^{(l+1)C_l} \mathscr{B}_l^{dC_l} M(\{C_l\})\,, \tag{7.12}$$

where

$$\mathscr{B}_l = \sum_{n=-\infty}^{\infty} \exp\left(-\frac{4\pi^2 \beta l t n^2}{L^2}\right). \tag{7.13}$$

Application of the Poisson summation formula yields

$$\mathcal{B}_l = L(4\pi\beta lt)^{-1/2} \sum_{n=-\infty}^{\infty} \exp\left(\frac{n^2L^2}{4\beta lt}\right). \tag{7.14}$$

If $4\pi\beta lt$ is larger than L, (7.13) converges faster than (7.14). Taking the limit $L\to\infty$ in (7.14) yields the well-known result for the infinite system (Feynman, 1976) but because we want to calculate the canonical partition function, the sum in (7.13, 14) is essential in order to get the correct answers. Experimentally we found that 2 to 4 of the largest terms in the series (7.13, 14) were sufficient to obtain high accuracy. Replacing the sum over permutations by a sum over cycles gives such a drastic reduction that it is possible to calculate the thermodynamic functions with a modest amount of computer time (seconds on a VAX 11/780). However this apparently efficient scheme to calculate the model properties in the canonical ensemble is not exactly the most convenient from a numerical point of view. Indeed, for fermions it does not solve the fundamental problem that one has to subtract two, possibly large numbers. To demonstrate the seriousness of this problem, we have plotted (see fig. 7.1) the number of digits that is lost in accuracy by adding the positive and negative contribution to the energy as a function of the temperature for the case of 8 (18) particles of a 4×4 (6×6) square lattice. It is clear that the number of lost digits increases with dimensionality and the number of particles in the system. Remark that the Monte Carlo method used for the 1D model does not suffer from the same problem because the sum over all configurations and permutations is performed in a completely different manner. In numerical work it is better to sum over all permutations before one changes the configuration.

It is interesting to examine the possibility of detecting the onset of boson condensation in a finite system. Of course we cannot expect that in the finite system there will be a true phase transition but it is good to know how a system of a finite number of non-interacting bosons mimics the infinite system. In

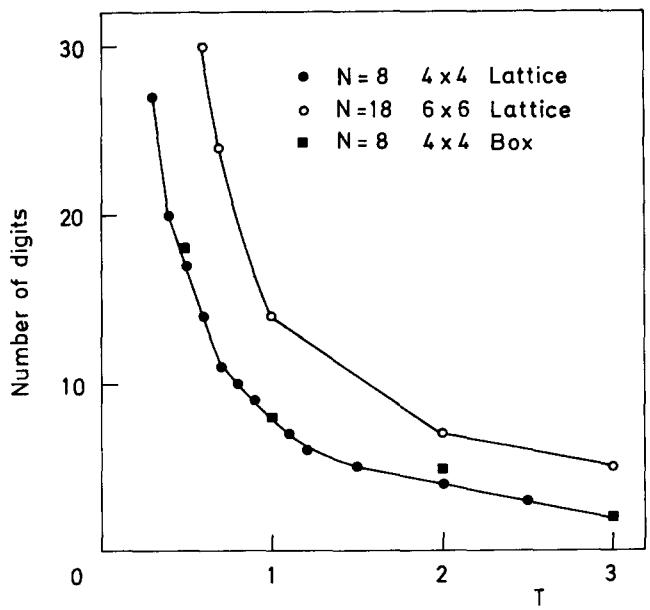

Fig. 7.1. Number of digits required to get one-digit accuracy of the fermion energy if one uses (7.10) to compute the thermodynamic functions, plotted as a function of the temperature T.

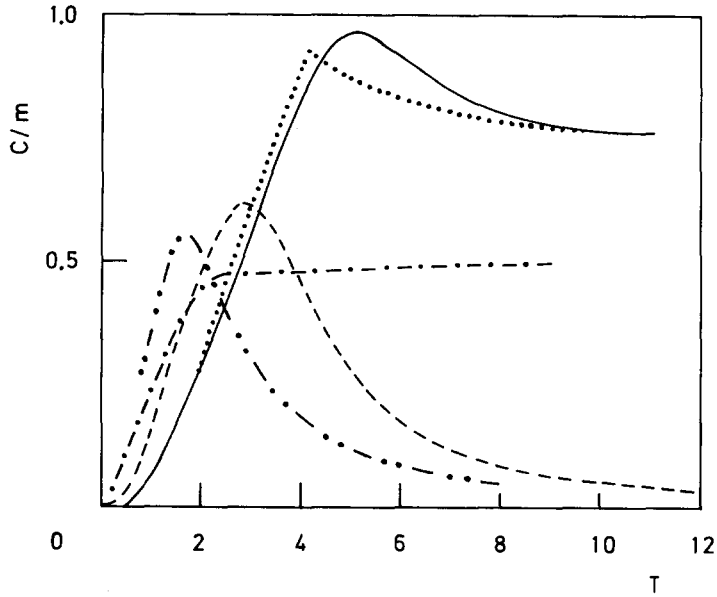

Fig. 7.2. Specific heat per site of a system of 18 non-interacting bosons on a 6×6 square lattice (dash-dot-dot) and in a square box (dash-dot) of size 6×6, subject to periodic boundary conditions, plotted as a function of the temperature T. Specific heat per site of a system of 13 non-interacting bosons on a $3 \times 3 \times 3$ cubic lattice (dashed line), in a cubic box (solid line) of size $3 \times 3 \times 3$ both with peridoic boundary conditions, and exact results obtained from the grand-canonical partition function of free bosons in a cubic box (dots).

fig. 7.2 we show results for the specific heat of the 2D bosons on a square lattice and a square box and for 3D bosons on a simple cubic lattice and a cubic box. For comparison the rigorous results for the specific heat of free bosons in a 3D box, obtained from the grand canonical partition function, are also shown. For the box models there is a notable difference between 2D and 3D. This could have been expected because in the 3D case Bose condensation occurs whereas in the 2D system nothing peculiar happens at finite temperature. Lattice models always exhibit a maximum in the specific heat. As the specific heat must vanish if $T \to 0$ and because the kinetic energy of particles hopping on a lattice is always bounded the specific heat should also vanish if $T \to \infty$.

8. Concluding remarks

We have discussed the application of the Monte Carlo method to problems of quantum statistical mechanics by considering explicit examples. As most Monte Carlo simulations for quantum systems at non-zero temperature have been performed within the context of the Trotter formula, we have paid considerable attention to the analysis of different Trotter-formula approximation schemes. Thereby lattice models, such as the ones treated in this review, offer the rather unique opportunity to compare with exact (numerical) results and make possible a thorough, quantitative evaluation of the efficiency and drawbacks of a particular approximation scheme.

For problems in classical statistical mechanics, implementation of the Monte Carlo algorithm is relatively simple. Implementation of a Monte Carlo algorithm for path-integral-like representations of the quantum models, constructed by means of the generalized Trotter formula, is not by any means straightforward. Simulations of spin-1/2 lattice systems require rather sophisticated and also dedicated

algorithms to deal with conservation laws. For one-dimensional lattice fermions this problem is not as acute as for the spin-1/2 model because the most important conservation law, particle conservation, can be built into the Monte Carlo procedure from the start. A general class of electron–phonon models (polarons) is ideally suited to be studied by the Monte Carlo technique. This is not very surprising since it is well-known that the path-integral approach is much more powerful than conventional Hamiltonian formulations when dealing with these systems.

The general conclusion is that combination of the Monte Carlo technique and the Trotter-formula approach to statistical mechanics can be very fruitful. However each new model that will (and should) be attacked, will require a lot of analytical as well as numerical groundwork before the actual Monte Carlo results can be trusted.

Acknowledgements

During this project we have benefitted a lot from discussions with Jan Fivez and Bart De Raedt. We are grateful to Victor V. Goldman for very helpful discussions on Bose condensation. The persistent advocating of the path-integral concept by Jozef T. Devreese has stimulated our interest in this field a lot. The authors thank Karl H. Michel for his continuous support and interest which were essential for our Monte Carlo work. We appreciate very much having received preprints from J.E. Hirsch, J.W. Lyklema, T. Schneider, M. Takahashi and M. Imada, and A. Wiesler. We gratefully acknowledge the continuous financial support by the project "Neutron scattering" of the Inter-University Institute for Nuclear Sciences (Belgium), the project "Supercomputers" of the National Fund for Scientific Research (Belgium) and the Dutch Stichting voor Fundamenteel Onderzoek der Materie. One of us (H.D.R.) thanks the National Fund for Scientific Research (Belgium) for financial support. We are grateful to the Stichting Academisch Rekencentrum Amsterdam for granting us about 140 hours of computer time on their CDC 170/750, and for giving us access to their CYBER 205 during its test phase.

References

Abramowitz, M. and I.A. Stegun, 1965, Handbook of Mathematical functions (Dover, N.Y.).
Appel, J., 1968, in: Solid State Physics, Vol. 21, eds. F. Seitz, D. Turnbull and H. Ehrenreich (Academic Press, N.Y.).
Barma, M. and B.S. Shastry, 1977, Phys. Lett. 61A, 15.
Barma, M. and B.S. Shastry, 1978, Phys. Rev. B18, 3351.
Baker Jr., G.A., G.S. Rushbrooke and H.E. Gilbert, 1964, Phys. Rev. A135, 1272.
Baxter, R.J., 1971, Phys. Rev. Lett. 26, 832.
Baxter, R.J., 1972a, Ann. Phys. (N.Y.) 70, 193.
Baxter, R.J., 1972b, Ann. Phys. (N.Y.) 70, 323.
Becker, W., B. Gerlach and H. Schliffke, 1983, Phys. Rev. B28, 5735.
Bethe, H.A., 1931, Z. Phys. 71, 205.
Betts, D.D., 1974, in: Phase Transitions and Critical Phenomena, Vol. 3, eds. C. Domb and M.S. Green (Academic, N.Y.)
Betts, D.D. and S.B. Kelland, 1983, J. Phys. Soc. Jpn. 52, 11.
Betts, D.D. and M. Plischke, 1976, Can. J. Phys. 54, 1553.
Betts, D.D., F.C. Salevsky and J. Rogiers, 1981, J. Phys. A14, 531.
Berezinskii, V.L., 1970, Zh. Eksp. Teor. Fiz. 59, 907 (Sov. Phys. JETP. 32 (1971) 493).
Binder, K., 1979, in: Monte Carlo Methods in Statistical Physics, ed. K. Binder (Springer, Berlin).
Bonner, J.C. and M.E. Fisher, 1964, Phys. Rev. A135, 640.
Bonner, J.C., H.W.J. Blöte, H. Beck and G. Müller, 1981, in: Physics in One Dimension, eds. J. Bernasconi and T. Schneider (Springer, Berlin).
Ceperley, D.M. and M.H. Kalos, 1979, in: Monte Carlo Methods in Statistical Physics, ed. K. Binder (Springer, Berlin).
Chakravarty, S. and D.B. Stein, 1982, Phys. Rev. Lett. 49, 582.
Cullen, J.J. and D.P. Landau, 1983, Phys. Rev. B27, 297.
Delbourgo, R. and P.D. Jarvis, 1983, Phys. Lett. 96A, 165.

Dekeyser, R., M. Reynaert, A.L. Stella and F. Toigo, 1978, Phys. Rev. B18, 3486.
De Raedt, B. and H. De Raedt, 1983, Phys. Rev. Lett. 50, 1926.
De Raedt, B. and H. De Raedt, 1983, Phys. Rev. B29, 5325.
De Raedt, H. and B. De Raedt, 1983, Phys. Rev. A28, 3575.
De Raedt, H., B. De Raedt, J. Fivez and A. Lagendijk, 1984, Phys. Lett. 104A, 430.
De Raedt, H., B. De Raedt and A. Lagendijk, 1984, Z. Phys. B 57, 209.
De Raedt, H. and A. Lagendijk, 1981, Phys. Rev. Lett. 46, 77.
De Raedt, H. and A. Lagendijk, 1982a, J. Stat. Phys. 27a, 731.
De Raedt, H. and A. Lagendijk, 1982b, Phys. Rev. Lett. 49, 1522.
De Raedt, H. and A. Lagendijk, 1983a, Phys. Rev. B27, 921.
De Raedt, H. and A. Lagendijk, 1983b, Phys. Rev. B27, 6097.
De Raedt, H. and A. Lagendijk, 1984, Phys. Rev. B30, 1671.
De Raedt, H., A. Lagendijk and J. Fivez, 1982, Z. Phys. B46, 261.
Des Cloizeau, J. and J.J. Pearson, 1962, Phys. Rev. 128, 2131.
De Sitter, J. and M. Goovaerts, 1973, Simon Stevin 46, 159.
Elliott, R.J., P. Pfeuty and C. Wood, 1970, Phys. Rev. Lett. 25, 443.
Emin, D., 1973, Adv. Phys. 22, 57.
Emin, D., 1982, Phys. Today, June, 34.
Emin, D. and T. Holstein, 1976, Phys. Rev. Lett. 36, 323.
Emery, V.J., 1979, in: Highly Conducting One-dimensional Solids, eds. J.T. Devreese, R.P. Evrard and V.E. van Doren (Plenum, N.Y.)
Faddeev, L.D., 1976, in: Methods in field theory, eds. R. Balian and J. Zinn-Justin (North-Holland, Amsterdam).
Ferdinand, A.E. and M.E. Fisher, 1969, Phys. Rev. 185, 832.
Feynman, R.P., 1955, Phys. Rev. 97, 660.
Feynman, R.P., 1972, Statistical Mechanics (Benjamin, Reading).
Feynman, R.P. and A.R. Hibbs, 1965, Quantum Mechanics and Path Integrals (McGraw-Hill, N.Y.).
Fisher, M.E. and A.E. Ferdinand, 1967, Phys. Rev. Lett. 19, 169.
Fosdick, L.D., 1963, J. Math. Phys. 3, 1251.
Fosdick, L.D. and H.F. Jordan, 1966, Phys. Rev. 143, 58.
Fowler, M. and M.W. Puga, 1978, Phys. Rev. B18, 421.
Fradkin, E. and J.E. Hirsch, 1983, Phys. Rev. B27, 1680.
Fröhlich, J. and T. Spencer, 1981, Phys. Rev. Lett. 46, 1006.
Fucito, F., E. Marinari, G. Parisi and C. Rebbi, 1981, Nucl. Phys. B180, 369.
Gaudin, M., 1971, Phys. Rev. Lett. 26, 1301.
Glimm, J. and A. Jaffe, 1981, Quantum Physics (Springer, N.Y.).
Golden, S., 1965, Phys. Rev. 137B, 1127.
Gutfreund, H. and M. Weger, 1977, Phys. Rev. B16, 1753.
Handscomb, D.C., 1962, Proc. Cambridge Philos. Soc. 58, 594.
Handscomb, D.C., 1964, Proc. Cambridge Philos. Soc. 60, 115.
Hamber, H.W., 1981, Phys. Rev. D24, 951.
Hammersley, J.M. and Handscomb, D.C., 1964, Monte Carlo Methods (Methuen, London).
Herman, M.F., E.J. Bruskin and B.J. Berne, 1982, J. Chem. Phys. 76, 5150.
Hirsch, J.E., 1983a, Phys. Rev. Lett. 51, 296.
Hirsch, J.E., 1983b, Phys. Rev. Lett. 51, 1900.
Hirsch, J.E., 1983c, private communication.
Hirsch, J.E. and E. Fradkin, 1982, Phys. Rev. Lett. 49, 402.
Hirsch, J.E. and E. Fradkin, 1983, Phys. Rev. B27, 4302.
Hirsch, J.E. and D.J. Scalapino, 1983a, Phys. Rev. Lett. 50, 1168.
Hirsch, J.E. and D.J. Scalapino, 1983b, Phys. Rev. B27, 7169.
Hirsch, J.E., D.J. Scalapino, R.L. Sugar and R. Blankenbecler, 1981, Phys. Rev. Lett. 47, 1628.
Hirsch, J.E., R.L. Sugar, D.J. Scalapino and R. Blankenbeckler, 1982, Phys. Rev. B26, 5033.
Holstein, T., 1959a, Ann. Phys. (N.Y.) 8, 325.
Holstein, T., 1959b, Ann. Phys. (N.Y.) 8, 343.
Hsue, C.S., K.Y. Lin and F.Y. Wu, 1975, Phys. Rev. B12, 429.
Huang, K., 1963, Statistical Mechanics (Wiley, N.Y.).
Hubbard, J., 1978, Phys. Rev. B17, 494.
Hulthén, L., 1938, Arkiv. Mat. Astron. Fysik 26A, No. 11.
Jackson, S.A. and P.M. Platzman, 1981, Phys. Rev. B24, 499.
Johnson, J.D., 1974, Phys. Rev. A9, 1743.
Johnson, J.D. and J.C. Bonner, 1980, Phys. Rev. Lett. 44, 616.
Johnson, J.D. and B.M. McCoy, 1972, Phys. Rev. A6, 1613.

Johnson, J.D., S. Krinsky and B.M. McCoy, 1973, Phys. Rev. A8, 2526.
Jordan, H.F. and L.D. Fosdick, 1968, Phys. Rev. 171, 128.
José, J.V., L.P. Kadanoff, S. Kirkpatrick and D.R. Nelson, 1977, Phys. Rev. B16, 1217.
Kalos, M.H., 1984, in: Monte Carlo Methods in Quantum Problems, ed. M.H. Kalos (Reidel, Dordrecht, Holland).
Kalos, M.H., D. Levesque and L. Verlet, 1974, Phys. Rev. A9, 2178.
Kagoshima, S., T. Ishiguro and H. Anzai, 1976, J. Phys. Soc. Jpn. 41, 2061.
Katsura, S., 1962, Phys. Rev. 127, 1508.
Kaufman, B., 1949, Phys. Rev. 76, 1232.
Kelland, S.B., D.D. Betts and J. Oitmaa, 1981, J. Phys. A14, 69.
Klauder, J.R., 1978, in: Path Integrals, eds. G.J. Papadopoulos and J.T. Devreese (Plenum, N.Y.).
Krinsky, S., 1972, Phys. Lett. 39A, 169.
Kogut, J.B., 1979, Rev. Mod. Phys. 51, 659.
Kondo, J. and K. Yamaji, 1977, J. Phys. Soc. Jpn. 43, 424.
Kosterlitz, J.M., 1974, J. Phys. C7, 1046.
Kosterlitz, J.M. and D.J. Thouless, 1973, J. Phys. C6, 1181.
Lagendijk, A. and H. De Raedt, 1982, Phys. Rev. Lett. 49, 602.
Landau, L. and E. Lifshitz, 1980, Statistical Physics, Third revised edition (Pergamon, Oxford).
Lang, C.B. and H. Nicolai, 1982, Nucl. Phys. B 200, 135.
Lieb, E.H. and D.C. Mattis, 1966, Mathematical Physics in One Dimension (Academic, N.Y.)
Lieb, E.H., T. Schultz and D.C. Mattis, 1961, Ann. Phys. (N.Y.) 16, 407.
Lieb, E.H. and W.E. Thirring, 1976, in: Studies in Mathematical Physics, eds. E.H. Lieb, B. Simon and A.S. Wightman (Princeton University Press, Princeton, N.J.).
Lieb, E.H. and F.Y. Wu, 1972, in: Phase Transitions and Critical Phenomena, Vol. 1, eds. C. Domb and M.S. Green (Academic, N.Y.).
Luther, A. and I. Peschel, 1975, Phys. Rev. B12, 3908.
Luttinger, J.M. and Chih-Yuan Lu, 1980, Phys. Rev. B21, 4251.
Lyklema, J.W., 1982, Phys. Rev. Lett. 49, 88.
Lyklema, J.W., 1983, Phys. Rev. B27, 3108.
Mahan, G.D., 1981, Many Particle Physics (Plenum, N.Y.).
Marinari, E., G. Parisi and C. Rebbi, 1981, Nucl. Phys. B190, 734.
Marland, L.G. and D.D. Betts, 1980, Phys. Lett. 76A, 271.
Martin, O. and S. Otto, 1982, Nucl. Phys. B203, 297.
Mattis, D.C., 1979, Phys. Rev. Lett. 42, 1503.
McMillan, W.L., 1965, Phys. Rev. A138, 442.
Mermin, N.D. and H. Wagner, 1966, Phys. Rev. Lett. 17, 1133.
Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, 1953, J. Chem. Phys. 21, 1087.
McCoy, B.M. and T.T. Wu, 1968, Nuovo Cimento B56, 311.
Müller, G., H. Beck and J.C. Bonner, 1979, Phys. Rev. Lett. 43, 75.
Müller, G., H. Thomas, H. Beck and J.C. Bonner, 1981, Phys. Rev. B24, 1429.
Oitmaa, J. and D.D. Betts, 1978, Can. J. Phys. 56, 897.
Onsager, L., 1944, Phys. Rev. 65, 117.
Ovchinnikov, A.A., 1972, Zh. Eksp. Teor. Fiz. 64, 342 (Sov. Phys. JETP. 37 (1973) 176).
Pearson, R.B., 1977, Phys. Rev. B16, 1109.
Peeters, F.M. and J.T. Devreese, 1982, Phys. Status Solidi B112, 219.
Pfeuty, P., 1970, Ann. Phys. (N.Y.) 57, 79.
Pfeuty, P. and R.J. Elliott, 1971, J. Phys. C4, 2370.
Pokrovsky, V.L. and G.V. Uimin, 1978, J. Phys. C11, 3535.
Pouget, J.P., S.K. Khanna, F. Denoyer and R. Comès, 1976, Phys. Rev. Lett. 37, 437.
Pouget, J.P., S.M. Shapiro, G. Shirane, A.F. Garito and A.J. Heeger, 1979, Phys. Rev. B19, 1792.
Reed, M. and B. Simon, 1972, Methods of Modern Mathematical Physics I (Academic, N.Y.).
Rogiers, J. and R. Dekeyser, 1976, Phys. Rev. B13, 4886.
Scalapino, D.J. and R.L. Sugar, 1981, Phys. Rev. B24, 4295.
Schneider, T. and E. Stoll, 1981, Phys. Rev. Lett. 47, 377.
Schulman, L.S., 1981, Techniques and Applications of Path Integration (Wiley, N.Y.).
Schweizer, J.S., R.M. Stratt, D. Chandler and P.G. Wolynes, 1981, J. Chem. Phys. 75, 1347.
Shoji, H. and N. Tokuda, 1981, J. Phys. C14, 1231.
Stanley, H.E. and T.A. Kaplan, 1966, Phys. Rev. Lett. 17, 913.
Stella, A.L. and F. Toigo, 1978, Phys. Rev. B17, 2343.
Su, W.P., J.R. Schrieffer and A.J. Heeger, 1980, Phys. Rev. B22, 2099.
Sumi, A. and Y. Toyozawa, 1973, J. Phys. Soc. Jpn. 35, 137.
Sutherland B., 1970, J. Math. Phys. 11, 3183.

Suzuki, M., 1966, J. Phys. Soc. Jpn. 11, 3183.
Suzuki, M., 1971, Phys. Lett. 34A, 94.
Suzuki, M., 1976a, Prog. Theor. Phys. 56, 1454.
Suzuki, M., 1976b, Commun. Math. Phys. 51, 183.
Suzuki, M., 1977, Commun. Math. Phys. 57, 193.
Suzuki, M., 1985, J. Math. Phys. 26, 601.
Suzuki, M., 1985, Phys. Rev. B31, 2957.
Suzuki M., S. Miyashita and A. Kuroda, 1977, Prog. Theor. Phys. 58, 1377.
Suzuki, M. and S. Miyashita, 1978, Can. J. Phys. 56, 904.
Symanzik, K., 1965, J. Math. Phys. 6, 1155.
Swendsen, R.H., 1982, Phys. Rev. Lett. 49, 1302.
Takahashi, M., 1973, Prog. Theor. Phys. 50, 1519.
Takahashi, M. and M. Imada, 1983, J. Phys. Soc. Jpn. 53, 963.
Takano, H. and M. Suzuki, 1981, J. Stat. Phys. 26, 635.
Tatsumi, T., 1981, Prog. Theor. Phys. 65, 451.
Thompson, C.J., 1965, J. Math. Phys. 6, 1812.
Tobochnik, J. and G.V. Chester, 1979, Phys. Rev. B20, 3761.
Trotter, H.F., 1959, Proc. Am. Math. Soc. 10, 545.
Yang, C.N. and C.P. Yang, 1966, Phys. Rev. 151, 258.
Villain, J., 1975, J. Phys. (Paris) 36, 581.
Wegner, F., 1967, Z. Phys. 206, 465.
Weingarten, D.H. and D.N. Petcher, 1981, Phys. Lett. 99B, 333.
Whitlock, P.A. and M.H. Kalos, 1979, J. Comp. Phys. 30, 361.
Wiegel, F.W., 1975, Phys. Reports 16C, 57.
Wiesler, A., 1982, Phys. Lett. 89A, 359.
Wilkinson, J.H., 1965, The Algebraic Eigenvalue Problem (Clarendon, Oxford).
Wilcox, R.M., 1967, J. Math. Phys. 8, 962.
Wilson, B.G. and P. Kumar, 1983, Phys. Rev. B27, 3076.
Witschel, W., 1975, J. Phys. A8, 143.
Wolf, D. and J. Zittartz, 1981, Z. Phys. B43, 173.
Zittartz, J., 1978, Z. Phys. B31, 63.
Zittartz, J., 1978, Z. Phys. B31, 79.
Zittartz, J., 1978, Z. Phys. B31, 89.

Note added in proof

After the first version of the manuscript was submitted, some papers on subjects treated in this review have been brought to our attention. Handscomb's method has been extended to the antiferromagnetic Heisenberg model (Lee, Joannopoulos and Negele, 1984). Exact numerical calculations for 1D $S = 1/2$ chains have been performed by Betsuyaka (Betsuyaka, 1984, 1985) and Morgenstern and Würtz (Morgenstern and Würtz, 1985). Satija et al. (Satija, Wysin and Bishop, 1985) have used Monte Carlo simulations to study the specific heat and susceptibility of a 1D easy-plane ferromagnet in an in-plane magnetic field. In a series of papers Marcu et al. (Marcu and Weisler, 1985; Marcu, Müller and Schmatzer, 1985) present high precision Monte Carlo calculations for $S = 1/2$ and $S = 1$ chains. A Monte Carlo method for quantum spin systems, based upon a coherent state representation has recently been proposed by Takano (Takano, 1985).

Betsuyaka, H., 1984, Phys. Rev. Lett. 53, 629.
Betsuyaka, H., 1985, Prog. Theor. Phys. 73, 319.
Marcu, M. and A. Wiesler, 1985, J. Phys. A (to appear).
Marcu, M., J. Müller and F.K. Schmatzer, J. Phys. A (to appear).
Morgenstern, I. and D. Würtz, 1985, Phys. Rev. B (to appear).
Lee, D.H., J.D. Joannopoulos and J.W. Negele, 1984, Phys. Rev. B 30, 1599.
Satija, I., G. Wysin and A.R. Bishop, 1985, Phys. Rev. B 31, 3105.
Takano, H., 1985, Prog. Theor. Phys. 73, 332.